

Boston College Law School
Digital Commons @ Boston College Law School

Boston College Law School Faculty Papers

2-25-2013

Carrots, Sticks, and Salience

Brian D. Galle

Boston College Law School, brian.galle@bc.edu

Follow this and additional works at: <http://lawdigitalcommons.bc.edu/lspf>



Part of the [Law and Economics Commons](#), and the [Tax Law Commons](#)

Recommended Citation

Brian D. Galle. "Carrots, Sticks, and Salience." *Tax Law Review* (2013).

This Article is brought to you for free and open access by Digital Commons @ Boston College Law School. It has been accepted for inclusion in Boston College Law School Faculty Papers by an authorized administrator of Digital Commons @ Boston College Law School. For more information, please contact nick.szydowski@bc.edu.

Carrots, Sticks, and Salience

BRIAN GALLE¹

I. INTRODUCTION

Money is usually a good motivator.² When market actors buy or sell things that have unfortunate consequences for others—negative externalities—a standard economic prescription is to raise the price of the externality-producing good.³ Alternatively, the bad actor can be paid to stop. For policymakers, choosing between these two options, the stick and the carrot, can be a difficult task even if people respond as expected to cash incentives.⁴ But sometimes human beings make mistakes. Social science shows that it is easy to overlook the tiny-print disclaimer that “price does not include \$4.95 shipping and handling” when ordering a shiny new set of knives.⁵ If people are sometimes similarly neglectful of the carrots and sticks policymakers deliberately offer to change their behavior, what happens to the policy? How should policymakers respond to that problem? And is it ever desira-

¹ Associate Professor, Boston College Law School. I am grateful for helpful comments from Ben Alarie, Neil Buchanan, Gerrit De Geest, David Gamage, Jacob Goldin, Louis Kaplow, Wojciech Kopczuk, Gregg Polsky, Alex Raskolnikov, Chris Sanchirico, Rich Schmalbeck, Darien Shanske, and Larry Zelenak, as well as attendees of presentations at Columbia Law School, Duke University Law School, and the Canadian Law and Economics Association.

² See, e.g., Sean Combs, *It's All About the Benjamins*, on No Way Out (Bad Boy Records 1997). But see *It's All About the Benjamins vs. Actually the Benjamins Are Only a Small Part of a Larger Set of Concerns*, *The Onion* (Sept. 1, 1999), <http://www.theonion.com/articles/its-all-about-the-benjamins-vs-actually-the-benjam,11548/>.

³ E.g., Jonathan Gruber, *Public Finance and Public Policy* 135 (3d ed. 2011) (discussing “corrective taxation” as a means of addressing negative externalities).

⁴ My own analysis of the carrot/stick problem can be found in Brian Galle, *The Tragedy of the Carrots: Economics and Politics in the Choice of Price Instruments*, 64 *Stan. L. Rev.* 797, 813-40 (2012). For earlier efforts, see Thomas Sterner, *Policy Instruments for Environmental and Natural Resource Management* 167-79 (2003); Howard F. Chang, *An Economic Analysis of Trade Measures to Protect the Global Environment*, 83 *Geo. L.J.* 2131, 2149-64 (1995); Giuseppe Dari-Mattiaci & Gerrit De Geest, *Carrots, Sticks, and the Multiplication Effect*, 26 *J.L. Econ. & Org.* 365, 369-75 (2010); Jonathan Baert Wiener, *Global Environmental Regulation: Instrument Choice in Legal Context*, 108 *Yale L.J.* 677, 701-96 (1999).

⁵ See Xavier Gabaix & David Laibson, *Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets*, 121 *Q.J. Econ.* 505, 505-23, 526-32 (2006); Vicki G. Morwitz, Eric A. Greenleaf, Edith Shalev & Eric J. Johnson, *The Price Does Not Include Additional Taxes, Fees, and Surcharges: A Review of Research on Partitioned Pricing* 1-23, 25-35, 38-40 (Feb. 26, 2009) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1350004.

ble deliberately to hide sticks or carrots? This Article considers these questions.

These questions are not wholly new, but the literature that grapples with them seriously is only a couple of years old. As recently as 2009, I asserted—at the time rightly, I believe—that it was generally agreed that taxes intended to change individual behavior should always be as visible as possible, or “salient,” in the marketplace.⁶ Thanks to important contributions by David Gamage and Darien Shanske, Hunt Allcott, Sendhil Mullainathan, and Dmitry Taubinsky, and Jacob Goldin, it is now becoming clearer that the analysis is not necessarily so straightforward.⁷ For that reason, in my earlier work on the choice between carrots and sticks, I deliberately left open the question whether my conclusions could hold up if market actors are sometimes imperfectly rational.⁸

I try here to advance these earlier works in a few different ways. For one, I want to refine the analysis offered so far. Gamage and Shanske argue that low-salience taxes can be just as effective at changing behavior if policymakers increase their price to make up for the taxes’ low visibility.⁹ Picking up that thread, I examine to what extent this claim is still true when externality producers vary in their ability to notice the change in price. I also consider to what extent it is ever optimal for the government deliberately to reduce or raise the salience of a stick. I conclude that often the ideal combination reduces salience somewhat, but without fully offsetting the behavioral effects of that reduction with increased prices. In essence, government must trade off four different considerations: the cost of under-incentivizing “naïve” individuals who cannot recognize the hidden stick, the cost of over-incentivizing “sophisticated” individuals who see the stick’s real price, the possible welfare losses from confusing consumers, and the possible benefits from bringing in more government revenues.¹⁰ Since these effects are rarely perfectly symmetrical, it sometimes will make

⁶ Brian Galle, *Hidden Taxes*, 87 Wash. U. L. Rev. 59, 78 n.89 (2009). I was not alone. See note 91.

⁷ Hunt Allcott, Sendhil Mullainathan & Dmitry Taubinsky, *Energy Policy with Externalities and Internalities*, 112 J. Pub. Econ. 72 (2014); David Gamage & Darien Shanske, *Three Essays on Tax Salience: Market Salience and Political Salience*, 65 Tax L. Rev. 19 (2011); Jacob Goldin, *Note, Sales Tax Not Included: Designing Commodity Taxes for Inattentive Consumers*, 122 Yale L.J. 258 (2012).

⁸ See Galle, note 4, at 849.

⁹ See Gamage & Shanske, note 7, at 71-74.

¹⁰ For the most part, my analysis omits possible distributive effects. Although I view distribution as a key question in the policy merits both of carrots and sticks, see Galle, note 4, at 817-20, and tax salience generally, see Galle, note 6, at 100-04, my analysis here already has so many moving parts, and we know so little about salience’s distributive consequences, that further discussion is best left for other work.

sense to give up some of one to improve things along one of the other dimensions.

Another contribution I will attempt to offer is to extend this “tax salience” discussion to price instrument theory.¹¹ So far, those two literatures have mostly not overlapped, which has led to some oversights. For instance, price instrument theory implies that there are potentially important variations between the two different ways that changes in price affect our decisions: relative cost, or substitution effects, and the total size of our budget, or income effects.¹² Differences in income effects are in my view the main justification for ever preferring carrots over sticks.¹³ Therefore, it is important to consider the impact of salience and offsetting adjustments on both substitution and income effects separately. To take one example, adjusting prices upwards to offset low salience could also shrink the budgets of affected consumers, which in turn may prevent consumers from buying goods that produce externalities. That is usually a good thing for negative externalities, but often a problem for positive externalities.

A second extension is to use the tax debate to help policymakers choose between carrots and sticks. A few commentators have noted that salience can also affect efforts to reward positive-externality producers, but other than cautioning that it might be possible to have too much of a good thing, these authors have not applied all of the formal analysis of price instrument theory.¹⁴ Here again I think some of my findings are surprising. For instance, in some cases it might be ideal to have carrots be totally hidden from view, at least from those who are likely to have been willing to produce the good without any encouragement. Indeed, it is possible that salience can improve carrots by such an extent that they might actually become a viable policy option, even though under standard price theory sticks are almost always the better choice.

¹¹ For overviews of the literature on price instruments, see Sterner, note 4, at 167–79; Cameron Hepburn, *Regulation by Prices, Quantities, or Both: A Review of Instrument Choice*, 22 *Oxford Rev. Econ. Pol’y* 226 passim (2006).

¹² See Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 *Harv. L. Rev.* 1089, 1095–96, 1120–21 (1972).

¹³ See Galle, note 4, at 838. For other arguments in favor of carrots, see Gerrit De Geest & Giuseppe Dari-Mattiacci, *The Rise of Carrots and the Decline of Sticks*, 80 *U. Chi. L. Rev.* 341, 367–69 (2013).

¹⁴ See Lilian V. Faulhaber, *The Hidden Limits of the Charitable Deduction: An Introduction to Hypersalience*, 92 *B.U. L. Rev.* 1307, 1339–44 (2012); Brian Galle & Jonathan Klick, *Recessions and the Social Safety Net: The Alternative Minimum Tax as a Countercyclical Fiscal Stabilizer*, 63 *Stan. L. Rev.* 187, 246 (2010); see also Brett M. Frischmann & Mark A. Lemley, *Spillovers*, 107 *Colum. L. Rev.* 257, 278 (2007) (noting, without further comment, that internalization theories of intellectual property assume that “a decisionmaker will rationally act to maximize her own welfare”).

It is worth emphasizing here at the beginning that in some ways this is a science fiction story. Most of the important questions about salience are still unanswered. It may turn out that price changes can only meaningfully be hidden when they are too small to have important policy effects. Or the initial salience of a price instrument may have important effects on how people learn about the instrument. My work here necessarily has to rely on supposition and the best available evidence.

It might also be argued that this project is not just science fiction, but actually the work of a mad scientist. Policy advocates (rightly, in my opinion) call for efforts to protect consumers from efforts to “shroud” prices.¹⁵ Whether manipulating salience is ethical, consistent with democratic values, or in the long run healthy for democracy is an important question.¹⁶ I largely set it aside here, though, because I argue that salience effects are likely inevitable, even if not deliberately invited by government actors. People need time to understand complex new policies, and the behavior of nongovernmental third parties can affect how the public perceives price instruments. Policymakers therefore often must confront the problem that the response to price instruments will depart from standard accounts of rational behavior; at a minimum, my analysis helps us to understand how policy should respond to those departures.

II. BACKGROUND

Before diving into the psychology of price instruments, some readers may find it helpful to have a brief review of what price instruments are and what economic theory thinks they are good for. Readers who know that story, but who have not encountered my earlier work on the subject, can skip to the following Section for a short recap of the best methods for choosing among different price instruments for different kinds of problems. Readers unfamiliar with recent developments in the tax salience literature may also wish to read Section C. All readers may also want to read that Section C briefly for an introduction to my terminology going forward, as well as some slight refinements to existing salience theory.

¹⁵ See Oren Bar-Gill & Elizabeth Warren, *Making Credit Safer*, 157 U. Pa. L. Rev. 1, 6-7 (2008); Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. Rev. 630, 721-49 (1999).

¹⁶ See Deborah H. Schenk, *Exploiting the Salience Bias in Designing Taxes*, 28 Yale J. on Reg. 253, 287-89 (2011) (arguing that manipulating salience can be consistent with democratic values).

A. *Price Instruments as a Response to Market Failure*

Modern economic theories of government regulation begin with the premise that markets sometimes fail.¹⁷ Externalities are a classic example.¹⁸ An externality, simply put, is a harm or benefit that affects someone other than the actor making an economic decision.¹⁹ When the poet Henry Wadsworth Longfellow “shot an arrow into the air, / It fell to earth [he] knew not where,”²⁰ that was an externality—and a particularly pointy one. Not all externalities are harmful or “negative.” When Newton noted that he stood “on ye shoulders of Giants,”²¹ he was acknowledging the positive externalities provided to him by great thinkers of the past.

Externalities are a problem in an unregulated market. Longfellow’s unmediated archery exercises are, at best, costly for neighbors who must invest in armor, or at least sturdy roofs. Absent some negotiated solution between Longfellow and these neighbors,²² Longfellow has no economic reason to care about his stray missiles or their in terrorem effects on neighboring investments. And clearly Newton has no way to negotiate with Brahe and Galileo to reward them for their efforts. True, most backyard archers have some conscience. And some inventors are motivated by glory, love of man, or professional norms.²³ But not all are, or in any event are not motivated enough to produce the amount of externality that would best meet society’s needs.

Economists offer a trio of standard solutions to the externalities problem. A first is regulation or prohibition, sometimes called “quantity regulation.”²⁴ “No more than one backyard arrow per day, or you go to jail.” Price instruments form a second broad category.²⁵ These include measures familiar from first-year law courses, such as tort liability, as well as taxes and subsidies.²⁶ The third group, disclosure and

¹⁷ Gruber, note 3, at 3.

¹⁸ Id. at 4.

¹⁹ Id. at 122.

²⁰ Henry Wadsworth Longfellow, *The Arrow and the Song*, in *Poems and Other Writings* 53 (J.D. McClatchy ed., Library of Am. 2000) (1845).

²¹ Letter from Isaac Newton to Robert Hooke (Feb. 5, 1675/6), in 1 *The Correspondence of Isaac Newton* 416, 416 (H.W. Turnbull ed., 1959).

²² As readers likely know, the possibility that parties may be able to bargain to achieve efficient outcomes itself has a long pedigree, most famously in R.H. Coase, *The Problem of Social Cost*, 3 J.L. & Econ. 1, 15 (1960).

²³ See Rebecca Tushnet, *Economies of Desire: Fair Use and Marketplace Assumptions*, 51 Wm. & Mary L. Rev. 513, 522-36 (2009).

²⁴ See Gruber, note 3, at 137.

²⁵ See Sterner, note 4, at 212-13.

²⁶ See De Geest & Dari-Mattiacci, note 13, at 343.

information sharing, either alone or in combination with other tools, can also help to move markets.²⁷

All three tools share a common goal of achieving what might be called the optimal level of externality.²⁸ Government could ban archery entirely, but then how will Longfellow hunt his dinner? More prosaically, eliminating even the worst pollutants is costly. Should government bankrupt coal producers, or is there a way to balance clean air against the costs of achieving it? On the positive externality side, everyone might agree that charity is beneficial. But how much should government spend to clothe or educate one more child?

Marginal analysis is the standard answer to these kinds of balancing questions.²⁹ Under this approach, the policymaker asks herself, "On the margin—that is, for the very next unit of good or bad produced—what is the harm or benefit of that one unit for *everyone in society*?" We might therefore call this the "marginal social damage," in the case of a negative externality, and "marginal social benefit" for a positive one. The policy maker then compares this harm or benefit against the marginal costs to the producer. If the producer's private marginal cost is greater than the marginal social damage, it does not pay, on net, to prevent the damage: Counting the producer's losses, society would lose by forcing the producer to avoid the externality.³⁰

This last point is a key one for later discussion, and deserves some extra emphasis. Assuming that we value the welfare of polluters as much as we value the well-being of the rest of the public (and, admittedly, that is a large assumption) then excessive reduction of pollution is just as bad as insufficient reduction.³¹ If *Pete Polluter* spends \$150 to avoid a harm that would cost his neighbors only \$100, society is worse off by \$50 overall. This \$50 overspending is usually called dead-weight loss—money or effort spent that accomplishes no other purpose.³² In other words, we have to count the costs of changing the quantity of an externality when we figure out how much of it we want.

²⁷ See Evert Vedung & Frans C.J. van der Doelen, *The Sermon: Information Programs in the Public Policy Process—Choice, Effects, and Evaluation*, in *Carrots, Sticks, and Sermons: Policy Instruments and Their Evaluation* 103, 103-28 (Marie-Louise Bemelmans-Videc, Ray C. Rist & Evert Verdung eds., 2003).

²⁸ See Gruber, note 3, at 137-39; Gloria E. Helfand, Peter Berck & Tim Maull, *The Theory of Pollution Policy*, in *1 Handbook of Environmental Economics* 249, 252-54 (Karl-Göran Mäler & Jeffrey R. Vincent eds., 2003).

²⁹ See Gruber, note 3, at 126.

³⁰ See *id.* at 124-25. Note, importantly, that for simplicity one assumes here that one should count the costs and benefits for the producer and everyone else equally. See *id.* That is a controversial proposition, but I leave it aside here for ease of exposition.

³¹ See *id.* at 139.

³² See *id.* at 51-52.

Price and quantity instruments differ in their approach to reaching this optimal level of externality.³³ Under a typical quantity regulation, the quantity the government should set, optimally, as the correct amount is the quantity at which private costs for further changes in the amount of externality produced first begin to exceed social benefits.³⁴ To do that, of course, the government must decide for itself the optimal level, which means that it must know not only the marginal social damage or marginal social benefit but also the private cost structure of producers.³⁵

Price instruments are typically a bit less demanding.³⁶ Suppose the government knows only the marginal social damage of an externality. It then can set a fee or tax—often called a Pigouvian tax, after its first expositor, the economist A.C. Pigou—equal to that amount.³⁷ Producers then reveal their cost structure to the government by deciding whether to pay the tax. Suppose the tax is \$100 and it costs *Wanda the Widget-Maker* \$101 to avoid creating widget sludge, a form of pollution. Wanda should go ahead and produce the sludge, paying the tax. That option is \$1 cheaper for her, and for society overall. By doing so,

³³ See *id.* at 137-46.

³⁴ See *id.* at 137. I am simplifying here for the sake of exposition. A more rigorous approach to setting the optimal quantity would also account for other factors that might affect the efficiency of the regulation. For example, if the regulation imposes costs and the expectation of those costs changes behaviors other than the production of the externality—for example, distorts consumer choices among products—the ideal regulation might balance disruption of these expectations against pollution control. See Helmuth Cremer, Firouz Gahvari & Norbert Ladoux, Externalities and Optimal Taxation, 70 J. Pub. Econ. 343, 346 (1998).

Another complication is that efficient regulation requires that the marginal costs of abating pollution be equal across all polluting sources—or, putting the same point a different way, government should ask for less abatement at firms where abatement is very expensive and more at those where it is cheap. See Robert N. Stavins, Environmental Economics, in 2 The New Palgrave Dictionary of Economics 886, 886-87 (Steven N. Durlauf & Lawrence E. Blume eds., 2d ed. 2008). Some simpler forms of regulation would have trouble satisfying this requirement, although the invention of tradable permits has made the task much easier for sources, such as those that do not produce “hot spots,” where trading is feasible. See *id.* at 887. Under certain assumptions, tradable permits are in many ways equivalent to a price instrument, see Hepburn, note 11, at 229, with the important exception that the permitting system allows for imposition of hard caps on the quantity. Thus quantity measures are usually thought to be preferable when it is important to get the quantity of externality right. See *id.* at 241, 243. But this may be inaccurate if governments can use tax schedules that vary together with the expected marginal harm or if government is stuck with a flat rate but can adjust it in response to new information. See Louis Kaplow & Steven Shavell, On the Superiority of Corrective Taxes to Quantity Regulation, 4 Am. L. & Econ. Rev. 1, 7-10 (2002).

³⁵ Gruber, note 3, at 140. For more discussion of the level of detail a government regulator might need to regulate optimally, see Jon D. Hanson & Kyle D. Logue, The Cost of Cigarettes: The Economic Case for Ex Post Incentive-Based Regulation, 107 Yale L.J. 1163, 1264-68 (1998).

³⁶ See Kaplow & Shavell, note 34, at 4.

³⁷ See A.C. Pigou, The Economics of Welfare 192-94 (4th ed. 1932).

she reveals to observers that her private costs are at least \$100. Alternately, government could use a carrot instead of a stick: It could offer to pay Wanda \$100 for each unit she reduces. Again, Wanda will accept if the cost for her of reducing sludge is less than \$100.

Because of this considerable informational advantage, a number of commentators favor price instruments, at least when everything else is equal.³⁸ Since my focus here is on how to design a price instrument, for now I do not elaborate on other factors that might influence government's choice between price and quantity regulation. Instead, I now consider a second choice—assuming regulators pick a price instrument, what *kind* of price instrument should they use?

B. The Tragedy of the Carrots: A Brief Review

Once policymakers decide to rely on a price instrument, they have a choice between rewarding or penalizing, between carrots and sticks. Both options have similar effects on the marginal incentives of externality producers.³⁹ Whether producers are rewarded, or nonproducers fined, compliance with the government's objective can save money relative to noncompliance.⁴⁰ The two mechanisms, however, vary in a number of other important ways. Which option is the better choice for a particular policy depends largely on these other factors.⁴¹

Sticks are, except in unusual circumstances, the more efficient tool for reigning in the social overproduction of some negative-externality-laden good.⁴² Sticks earn the government money, while carrots drain the treasury, wasting hard-won tax revenues.⁴³ Revenue is critical because raising taxes is costly: In addition to paying the tax, many people also change their behavior to minimize taxes, causing deadweight loss.⁴⁴ In addition, carrots give producers more resources to create the unwanted good.⁴⁵ Similarly, in many cases, as individuals get wealthier, they demand more of the undesirable product, a phenomenon known as the income effect.⁴⁶ Carrots are also wasteful if producers plan to cut back on their activities anyway. And overproducers who know they will be paid to curtail their activities in the future have

³⁸ See, e.g., Hepburn, note 11, at 228-29.

³⁹ See Helfand et al., note 28, at 277-78.

⁴⁰ See id. at 278.

⁴¹ See Galle, note 4, at 809-13.

⁴² For development of the points in this paragraph, see id. at 813-31.

⁴³ For a caveat to this point, see text accompanying notes 112-15.

⁴⁴ See Gruber, note 3, at 51-52.

⁴⁵ See id. at 36.

⁴⁶ See id. at 36. For example, poorer commuters may take the bus, while richer ones may prefer to drive. See id.

an incentive to begin overproducing, while the opposite is true of sticks.

In contrast, carrots are more defensible for encouraging the production of a good with positive externalities, where we would expect social underproduction.⁴⁷ In that case, the fact that carrot recipients have more resources is desirable, since we want them to produce or demand more of the good.⁴⁸ On the other hand, it is still the case that the expectation of future carrots has unwanted incentive effects, encouraging producers to delay producing the good until the government agrees to pay them. And carrots remain costlier, especially when factoring in the possibility that some might altruistically produce the good without subsidy. So although carrots are less clearly dominated by sticks in the positive externality setting, there remains a question whether they are worth the cost.

To see the important difference that income effects may have on the production of externalities, consider Figure 1. This figure graphs, in a highly simplified way, the private costs of reducing (increasing) negative (positive) externalities against their private social costs (benefits). Government typically will set the price of a carrot or stick at the intersection of these two curves, which one can call the equilibrium price, represented here by point A. Once more, this is the optimal price: At any higher point, the tax will wastefully overincentivize producers, while at any lower point it will leave on the table some additional, efficient, changes in externality levels. The rightward arrow depicts the impact of carrots on the production of positive externalities, or the effect of sticks on the curtailment of negative externalities.⁴⁹ By increasing the wealth of, say, donors to charity, government has caused each donor to be willing to create more charity for any given dollar amount of subsidy. Thus, the income effect of the subsidy increases total output, without affecting the efficient subsidy amount. Sticks for positive externalities, or carrots for negative would move the cost curve leftwards, diminishing the efficacy of government incentives.⁵⁰

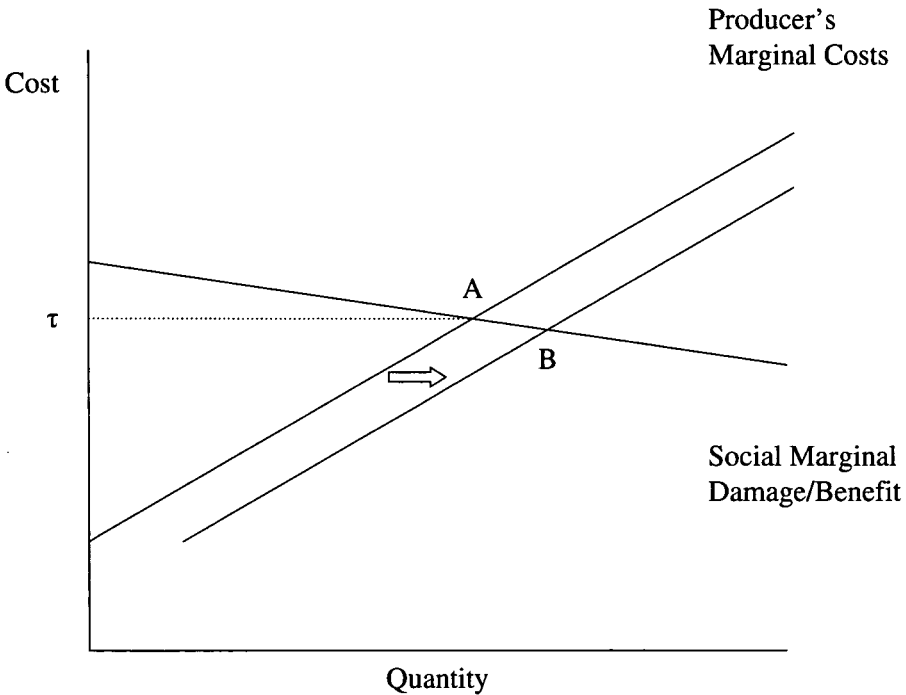
⁴⁷ See *id.* at 43-50 (noting that unregulated markets tend to underproduce goods with positive externalities).

⁴⁸ See Galle, note 4, at 832.

⁴⁹ Sticks can shift the private cost curve rightwards for two reasons. In the case of individuals who pay the stick, their demand for the polluting good declines. They will demand less of it for each dollar of government incentive. In the case of firms, the stick reduces output, meaning that there are fewer polluting goods sold.

⁵⁰ Information techniques can function similarly; by allowing firms to discover best practices for, say, clean production processes, the government encourages firms to reduce their private costs, shifting the curve rightwards. See Lori Snyder Benneer & Robert N. Stavins, *Second-Best Theory and the Use of Multiple Policy Instruments*, 37 *Envtl. & Resource Econ.* 111, 121 (2007).

FIGURE 1
INCOME EFFECTS AND PRICE INSTRUMENTS



Despite the theoretical inferiority of carrots in most situations, they are extremely popular as policy tools.⁵¹ I argue that this unfortunate result is the product both of straightforward public choice considerations as well as less-obvious aspects of U.S. law and government structure.⁵² For example, decentralization encourages carrots, because sticks usually redistribute away from externality producers, and competing subnational governments typically cannot engage in extensive redistribution.

It is worth mentioning here two potential questions about my analysis that I did not consider in my earlier work. Kaplow might argue that most of the differences between the two instruments would disappear if either were enacted together with a perfectly offsetting tax or tax cut.⁵³ I do not disagree, and even agree that such perfect offsets

⁵¹ See, e.g., Hepburn, note 11, at 236-37; Frans L. Leeuw, *The Carrot: Subsidies as a Tool of Government—Theory and Practice*, in *Carrots, Sticks, and Sermons*, note 27, at 77, 77 (reporting that major European countries spend between 20% and 35% of their GDP on grants and subsidies); cf. Calabresi & Melamed, note 12, at 1117 (noting that their rule four, in which victims compensate polluters for limiting their right to pollute, “may well be the most frequent device employed”).

⁵² Galle, note 4, at 840-45.

⁵³ See Louis Kaplow, *The Theory of Taxation and Public Economics* 13-35 (2008), for general discussion, and Louis Kaplow, *Taxes, Permits, and Climate Change*, in *U.S. Energy*

might often be theoretically ideal. My goal is only to consider the second-best outcomes in the absence of optimal offsets. That is, I analyze the implementation of the price instrument in isolation from any such offsets, which after all so far have not been observed in practice. In addition, in the context I analyze here, where the incidence of a Pigouvian tax depends on the foibles of human cognition, enacting perfectly offsetting income tax adjustments may well be infeasible.

There may also be some instances in which income effects do not have the impact I have described so far. In some cases, reducing a negative externality could involve *increasing* demand from externality producers. For example, families might be encouraged to switch out old air conditioners that leak Freon for newer models. But curtailing negative externalities most commonly involves reducing the undesirable good, such as when we want to contain pollution, smoking, or fatty foods.⁵⁴ And when government wants to develop positive externality production, it is usually attempting to spur the creation of new goods, such as hospitals or research labs.⁵⁵ These are the archetypes on which my claims about income effects rest.

C. *Three Dimensions of Salience*

Until very recently the modern literature on price instruments relied on the assumption that economic actors respond rationally to price changes.⁵⁶ As I noted at the outset, though, evidence increasingly suggests that this assumption is implausible. For example, many studies now document consumers' failures to maximize what appears to outside observers to be the consumers' subjective well-being.⁵⁷ Consumers apparently underweight "hidden" fees, such as ongoing small service charges or extra markups for shipping or Internet access.⁵⁸

Tax Policy 168, 186-88 (Gilbert E. Metcalf ed., 2010), for analysis specifically in the carbon tax context. I am grateful to Kaplow for pressing this point in conversation.

⁵⁴ See Galle, note 4, at 816-17.

⁵⁵ See *id.* at 838-39.

⁵⁶ See Hepburn, note 11, at 235.

⁵⁷ For reviews, see B. Douglas Bernheim & Antonio Rangel, Behavioral Public Economics: Welfare and Policy Analysis with Nonstandard Decision Makers, in Behavioral Economics and Its Applications 7, 10-65 (Peter Diamond & Hannu Vartiainen eds., 2008); Stefano DellaVigna, Psychology and Economics: Evidence from the Field, 47 J. Econ. Literature 315 *passim* (2009).

⁵⁸ Hyeong Min Kim & Luke Kachersky, Dimensions of Price Salience: A Conceptual Framework for Perceptions of Multi-Dimensional Prices, 15 J. Product & Brand Mgmt. 139, 139-47 (2006); Aradhna Krishna, Richard Briesch, Donald R. Lehmann & Hong Yunn, A Meta-Analysis of the Impact of Price Presentation on Perceived Savings, 78 J. Retailing 101, 101-18 (2002); Vicki G. Morwitz, Eric A. Greenleaf & Eric J. Johnson, Divide and Prosper: Consumers' Reactions to Partitioned Prices, 35 J. Marketing Res. 453, 453-63 (1998).

The literature on tax salience supplies the most direct evidence that price instruments may not function as predicted by classic theory. Recent real-world and laboratory results find that taxpayers appear to neglect at least some of the economic consequences of taxes when they make decisions.⁵⁹ Consumers preferred household products whose posted prices did not include the sales tax to those where the tax was included,⁶⁰ were less aware of road tolls when their toll was debited automatically from an electronic account,⁶¹ and bought more cigarettes when taxes were not included in the posted price.⁶²

Taxpayers are not necessarily totally ignorant of low-salience taxes. Raj Chetty, Adam Looney, and Kory Kroft speculate that taxpayers may be aware that there is tax, but simply choose not to calculate its precise impact.⁶³ Alternately, as I have argued, taxpayers may be aware of the tax but lack the cognitive ability or willpower to make exact calculations.⁶⁴ In line with these theories, Naomi Feldman and Bradley Ruffle find in experiments that their subjects tend to respond *partially* to the full price of a tax.⁶⁵ Similarly, Jacob Goldin models salience as a discount factor, θ .⁶⁶ When θ is one, the taxpayer accurately computes the cost of the tax, but when θ is below one, she ignores some or all of the cost of the tax.

In some instances taxpayers may actually overestimate the cost of a tax. For example, Benjamin Miller and Kevin Mumford found that real households overestimated the benefits they would gain from the 2003 expansion of the Child and Dependent Care Credit.⁶⁷ Because of complex interactions with other tax provisions, the value of that credit was less than its “sticker” price for many families, but it appears that families responded to the credit’s incentive to hire paid childcare assistance as though they all were getting its full value.⁶⁸ Similarly, James Sallee reports that the notoriously complicated Alternative

⁵⁹ For reviews, see Gamage & Shanske, note 7, at 26-54, and Allcott et al., note 7, at 79.

⁶⁰ Raj Chetty, Adam Looney & Kory Kroft, *Salience and Taxation: Theory and Evidence*, 99 *Am. Econ. Rev.* 1145, 1153-56 (2009).

⁶¹ Amy Finkelstein, *EZ-Tax: Tax Salience and Tax Rates*, 124 *Q.J. Econ.* 969, 980-83 (2009).

⁶² Jacob Goldin & Tatiana Homonoff, *Smoke Gets in Your Eyes: Cigarette Tax Salience and Regressivity*, 5 *Am. Econ. J.: Econ. Pol’y* 302, 328-31 (2012).

⁶³ See Chetty et al., note 60, at 1173-75.

⁶⁴ See Galle, note 6, at 81-85.

⁶⁵ See Naomi E. Feldman & Bradley J. Ruffle, *The Impact of Tax Exclusive and Inclusive Prices on Demand* 25 (Aug. 2012) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2026937.

⁶⁶ Goldin, note 7, at 273.

⁶⁷ See Benjamin M. Miller & Kevin J. Mumford, *Personal Income Tax Salience: Evidence from the Child and Dependent Care Credit Expansion* 2 Krannert School of Mgmt., Purdue U., Working Paper No. 1261, 2014), available at www.Krannert.purdue.edu/faculty/kjmmfor/papers/Personal%20Income%20Tax%20Salience.pdf.

⁶⁸ See *id.* at 8-17.

Minimum Tax left some hybrid-car purchasers “surprised to find that they could not benefit from the [hybrid vehicle] credit.”⁶⁹ That is, buyers bought their cars expecting to get a tax benefit, and then did not. Likewise Jacob Goldin and Yair Listokin describe their online survey in which some participants erroneously believed they were benefitting from some charitable contribution deductions and the home mortgage interest deduction.⁷⁰ And, finally, Shanjun Li, Joshua Linn, and Erich Muehlegger find that gas buyers actually seem more responsive to changes in gas tax than to other changes in the costs of fuel, which they suggest could be related to more extensive media and consumer attention.⁷¹ In Goldin’s terms, all these taxpayers appear to have θ greater than one.

Lilian Faulhaber chooses a less mathematical nomenclature, dubbing these overestimates of tax price “hypersalience,” and suggests two possible explanations.⁷² One explanation is simply complexity.⁷³ What looks like a benefit is not always a benefit. For example, as Lawrence Zelenak argues, taxpayers who do not understand the tax code’s complex system of phase-outs may believe there is a tax incentive for them when there is not.⁷⁴

Third parties, especially those who might capture some of the benefit of the price instrument, can be another source of hypersalience. Faulhaber’s key example is fundraising by nonprofits: Since charities benefit when donors want to give more, it is in the charities’ interests to make taxpayers highly aware of the benefits of giving.⁷⁵ It is rather less in the charities’ short-term interest to make donors aware of significant limits on potential tax benefits, such as the annual cap on deductible donations.⁷⁶ Other examples abound. Hardware stores, not usually a good source of tax advice, were in 2010 plastered with signs touting “weatherization tax credits.”⁷⁷ Kelly Gallagher and Erich

⁶⁹ James M. Sallee, *The Surprising Incidence of Tax Credits for the Toyota Prius*, 3 *Am. Econ. J.: Econ. Pol’y* 189, 216 (2011).

⁷⁰ Jacob Goldin & Yair Listokin, *Tax Expenditure Salience*, 15 *Am. L. & Econ. Rev.* __ (forthcoming 2014), (manuscript at 5-8), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2097836.

⁷¹ See Shanjun Li, Joshua Linn & Erich Muehlegger, *Gasoline Taxes and Consumer Behavior* 6 *Am. Econ. J.: Econ. Pol’y* (forthcoming 2014) (manuscript at 12-13), available at <http://www.nber.org/papers/w17891>.

⁷² Faulhaber, note 14, at 1316-18.

⁷³ *Id.* at 1317.

⁷⁴ See Lawrence Zelenak, *Complex Tax Legislation in the TurboTax Era*, 1 *Colum. J. Tax L.* 91, 110-11 (2010).

⁷⁵ Faulhaber, note 14, at 1328-30.

⁷⁶ See *id.* at 1329.

⁷⁷ See, e.g., Home Depot, *Get Credit for Saving Energy: 6 Weeks Left for Eligible 2011 Tax Credit Purchases*, <http://community.homedepot.com/t5/Install-Replace/Get-Credit-for-saving-energy-6-weeks-left-for-eligible-2011-Tax-/td-p/33807>.

Muehlegger claim to find evidence along these lines, arguing that their finding of greater impact of hybrid-car tax credits around tax season implies an important role for advertising and salience more generally.⁷⁸

In addition to affecting the decision to buy or sell goods and services, salience can also affect voters' political preferences for various tax regimes. As I have observed, and as Deborah Schenk and Gamage and Shanske emphasize, a tax may be highly salient for market purposes but largely invisible politically, or vice versa.⁷⁹ Convincing evidence that "political salience,"⁸⁰ actually changes tax rates is considerably scarcer than the evidence for "market" salience.⁸¹ The closest is likely Marika Cabral and Caroline Hoxby, who report that property taxes are higher in jurisdictions where a larger portion of homeowners do not pay their property tax bills directly.⁸² Salience may also affect other interactions with the government, such as the decision to file legal claims or challenge tax assessments.⁸³

The literature also identifies but so far has not emphasized a possible difference in the saliences of income and substitution effects. Chetty, Looney, and Kroft explain that a consumer who neglects the impact of a tax potentially is making two different errors: One is a mistake of comparison shopping, the other a mistake of budgeting.⁸⁴ Taxes matter for relative prices, or substitution effects, such as when we decide whether paying a bit more for bouncier curls is worthwhile.⁸⁵ But if they add up to real money, taxes should also matter in terms of which goods buyers should want: Can we afford hair product at all, or should we be saving up for rent? This is the income effect. Chetty, Looney, and Kroft suggest that tax salience may have a differ-

⁷⁸ See Kelly Sims Gallagher & Erich Muehlegger, *Giving Green to Get Green: Incentives and Consumer Adoption of Hybrid Vehicle Technology*, 61 *J. Envtl. Econ. & Mgmt.* 1, 2 (2011).

⁷⁹ See Galle, note 6, at 99; Gamage & Shanske, note 7, at 54-59; Schenk, note 17, at 272-85.

⁸⁰ Gamage & Shanske, note 7, at 33-35; Schenk, note 16, at 272-74.

⁸¹ See Gamage & Shanske, note 7, at 34-54.

⁸² See Marika Cabral & Caroline Hoxby, *The Hated Property Tax: Salience, Tax Rates, and Tax Revolts* 37 (Nat'l Bureau Econ. Research, Working Paper No. 18514, 2012), available at <http://www.nber.org/papers/w18514>. Another suggestive result is Asmus Leth Olsen, *The Politics of Digits: Evidence of Odd Taxation*, 154 *Pub. Choice* 59, 67-71 (2013), who finds that the last digits of Danish municipal tax rates tend to be closer to nine than one, perhaps implying that officials, like merchants, emphasize the leftmost digit of a price.

⁸³ See Andrew Hayashi, *The Legal Salience of Taxation* 31-37 (Sept. 14, 2012) (unpublished manuscript), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2151867.

⁸⁴ See Chetty et al., note 60, at 1173-74.

⁸⁵ The author no longer has this problem. See Boston College Law School, <http://www.bc.edu/content/bc/schools/law/fac-staff/deans-faculty/galleb.html> (demonstrating author's advanced male-pattern baldness).

ent impact on income effects than on substitution effects, depending on how consumers budget.⁸⁶ For example, a household planner might look at last month's bank balance to decide how to shop for the coming months; even if the planner does not know that it was taxes that ate up part of her savings, she knows what it cost to live the way she did. The tax would then be salient for income-effect purposes but not for substitution effects.

The salience of income effects has mostly been neglected in the literature so far but is potentially fairly important. The split between income and substitution effects is a key factor in the choice of price instrument. In some cases, I argue, different salience for the two effects either undermines or allows for a more effective design of price instruments.⁸⁷

III. SALIENCE AND INSTRUMENT PRICING

What happens to the standard analysis of price instruments, and the nuances I attempted to add, when the targets of those instruments are inattentive or overly attentive to their effects? To help keep things simple, I focus my analysis on two examples of price instruments. The first is a carbon tax imposed on consumer purchases of carbon-producing goods. The second is the charitable contribution deduction, a benefit granted by federal tax law⁸⁸, as well as a number of states and other countries⁸⁹ to encourage donations to charity. Since the deduction reduces an expected cost, it can be thought of as a carrot. "Charity" is typically defined by law as an endeavor that produces goods with beneficial spillovers for people other than the purchaser, such as schools and arts institutions.⁹⁰ The analysis thus considers two archetypical price instruments, one a stick aimed at curbing negative externalities and the other a carrot to tempt producers of positive-externality goods.

⁸⁶ Chetty et al., note 60, at 1173-74.

⁸⁷ More generally, if salience reduces income effects, then low-income-salience taxes would be relatively less efficient than the salience literature so far predicts. Income effects are usually thought to help counterbalance the tendency of taxes to discourage work. I have previously argued that low-salience taxes could allow for more progressive taxation by diminishing this disincentive. See Galle, note 6, at 104-05. But I now partially repent of that view: If low-substitution-salience is combined with low-income-salience, the combined impact of the salience could potentially on net reduce incentives to work.

⁸⁸ IRC § 170.

⁸⁹ See Roger Colinvaux, Brian Galle & Eugene Steuerle, Evaluating the Charitable Deduction and Proposed Reforms 6 (Urban Inst., June 2012), available at <http://www.urban.org/UploadedPDF/412586-Evaluating-the-Charitable-Deduction-and-Proposed-Reforms.pdf>.

⁹⁰ E.g., Bruce R. Hopkins, *The Law of Tax-Exempt Organizations* 155 (10th ed. 2011).

I argue that the potential implications of salience are pretty wide-ranging, but cluster around two big questions. First, does salience affect the price the government should set for the instrument? And second, does salience affect which instrument is the right tool for a particular problem? I discuss the first in this Part. To be more precise, I examine whether government should alter the value of price instruments to account for salience, whether it should alter salience to account for the desired impact of the instrument, or maybe a little bit of both.

In fact, my central claim in this Part is that the government's best strategy will often be to mix price adjustments with adjustments to salience itself. Prior analyses of the salience of Pigouvian taxes have argued for or against fully offsetting the impact of salience with changes in price. But I try to show that with plausible assumptions neither of those choices is ideal. I then extend that basic result to a number of permutations—testing whether we would want different outcomes if we switch from carbon tax to the charitable contribution, from low-salience to hypersalient instruments, and if there is correlation between an externality producer's propensity to recognize the true price of the instrument and their costs of complying with the government's preference.

A. *The Basic Case: Carbon Taxes and Low Salience*

Most early commentators on tax salience assumed that Pigouvian taxes should always be fully salient,⁹¹ but it is now clear that story was at best incomplete. Two recent articles, by Jacob Goldin, and Gamage and Shanske, respectively, argue that a low-salience tax on externalities might outperform a fully salient tax. The articles differ in that Goldin defends the possible efficiency of a low-salience tax standing alone,⁹² while Gamage and Shanske argue for the efficiency of a low-salience tax whose price is adjusted to account for its salience.⁹³ I think both are right but that there is a third outcome potentially better than either. To see why, I need to unpack their arguments some more. For readers most interested in math, I also offer some relatively informal modeling—assuming for simplicity a world in which cost curves are linear over the range of interest—in the margins.

⁹¹ See Austan Goolsbee, *The TurboTax Revolution: Can Technology Solve Tax Complexity?*, in *The Crisis in Tax Administration* 124, 138 (Henry J. Aaron & Joel Slemrod eds., 2004); Chetty et al., note 60, at 1173; Finkelstein, note 61, at 981; Galle, note 6, at 80; Jacob Nussim, *To Confuse and Protect: Taxes and Consumer Protection*, 1 *Colum. J. Tax L.* 218, 244-47 (2010); Schenk, note 16, at 276 n.101.

⁹² Goldin, note 7, at 286.

⁹³ Gamage & Shanske, note 7, at 71.

1. *Low Salience Sticks, No Price Adjustment*

For Goldin, revenue is the key to the possible efficiency of a low-salience stick, even if its price is not adjusted to account for salience effects.⁹⁴ Recall that raising revenue through sticks can be less distortionary than many other forms of tax. Take the carbon tax example. Imagine a carbon tax priced so that each consumer paid exactly the marginal social damage she inflicted by consuming each unit of carbon. This money is then used to replace the corporate income tax. The corporate income tax is a highly distortive tax—that is, it produces a lot of deadweight loss because it changes many forms of behavior, ranging from entities' decisions about when and how to do business to the labor supply of the workers whose salaries may be impacted by it.⁹⁵ If governments replace a highly inefficient revenue source, such as the corporate income tax, with sticks, their economies should expand, a result that is well-known in the carbon tax literature.⁹⁶

⁹⁴ Goldin, note 7, at 286-87. In contrast, Allcott, Mullainathan, and Taubinsky explicitly omit revenue considerations, and so reach considerably different results. Allcott et al., note 7, at 10 & n.5.

⁹⁵ For extended discussion, see Alan J. Auerbach, *Taxation and Corporate Financial Policy*, in 3 *Handbook of Public Economics* 1251, 1266-87 (Alan J. Auerbach & Martin Feldstein eds., 2002).

⁹⁶ The environmental economics literature now suggests that carbon taxes may not produce this kind of efficiency gain if they are used to replace income taxes or other kinds of consumption taxes. Ian W.H. Parry & Wallace E. Oates, *Policy Analysis in the Presence of Distorting Taxes*, 19 *J. Pol'y Analysis & Mgmt.* 603, 604-10 (2000), provides an accessible overview; for a more technical discussion, see A. Lans Bovenberg & Lawrence H. Goulder, *Environmental Taxation and Regulation*, in 3 *Handbook of Public Economics*, note 95, at 1471, 1483-513. In essence, critics of this "revenue recycling" or "double dividend" argument say that the distortionary effect of a carbon tax is equivalent to that of any other sales tax. By reducing the bundle of goods workers can buy with their labor, it may affect workers' decisions about how hard to work. It also may change their decision about which product to buy from the more-polluting to the less-polluting good. To be sure, this change is on net socially desirable, but we still must account for the fact that the buyer's own preferences have been foiled, which is a form of deadweight loss. Still, even the strongest critics appear to concede that externalities aside, the carbon tax is on net less distortive than corporate income taxes. Cf. Bovenberg & Goulder, *supra*, at 1507 (claiming that environmental taxes may be more efficient than taxes on capital income). Since every important polluting nation has a corporate income tax, these opportunities for what I call revenue gains should be widely available. See Bennear & Stavins, note 50, at 116-17 (stating that the combination of a Pigouvian tax with a reduction in the most-distortive other tax sources is the most efficient policy).

Alternately, governments may make additional investments with their new stick revenues. If governments were perfectly rational investors, the use of a relatively inefficient tax might imply that government had some quite valuable investment opportunities available. That is, if government is willing to use a very costly revenue source for investing in public projects, those projects must have a very high payoff. Access to a new source of much cheaper revenue implies that the government could make profitable additional investments at that high payoff rate, or at least in a series of investments between that point and the point corresponding to the cost of stick funds on their investment-returns curve.

It might also be argued in defense of Goldin's position that hidden sticks have a lesser effect on labor/leisure decisions than the revenues they replace. Economists generally assume that consumption taxes can discourage work, on the theory that what motivates us to get out of bed is the purchasing power our labor buys us.⁹⁷ Consumption taxes (or other changes in the price of a product, such as a stick) reduce purchasing power and therefore diminish the incentive to work. Price instruments with a low market salience may diminish this effect—or, conceivably, some instruments could have full salience when it comes to the choice among different consumption items, but have reduced salience as to the labor/leisure decision.⁹⁸ One could call this extra dimension “labor/leisure salience,” although it is really still just a subspecies of market salience.

Low-salience sticks can be efficient to the extent that they facilitate this exchange of inefficient tax revenue for more-efficient stick revenue.⁹⁹ Externality producers who fail to notice the stick still must pay it. For example, assume that the marginal social damage of widget consumption is \$100, and that the corresponding carbon tax on wid-

But of course public choice and similar economic analyses of government decisionmaking suggest that the rational investor model likely does not capture how governments plan.

⁹⁷ See, e.g., Kaplow, note 53, at 55-56, 80-81. Differences in administrative and compliance costs may drive a wedge between the behavioral responses to income and consumption taxes, however. See Joel Slemrod, *Does It Matter Who Writes the Check to the Government? The Economics of Tax Remittance*, 61 Nat'l Tax J. 251 passim (2008).

⁹⁸ To suggest one causal mechanism for this difference, suppose that it does not occur to workers that consumption taxes should affect their work incentives. For evidence, see Tomer Blumkin, Bradley J. Ruffle & Yosef Ganun, *Are Income and Consumption Taxes Ever Really Equivalent? Evidence from a Real-Effort Experiment with Real Goods*, 56 Eur. Econ. Rev. 1200, 1206 (2012). Another possibility is that cognitive effort has different effects on sellers of labor and purchasers of goods. See Andrew Hayashi, Brent K. Nakamura & David Gamage, *Experimental Evidence of Tax Salience and the Labor-Leisure Decision: Anchoring, Tax Aversion, or Complexity?*, 41 Pub. Fin. Rev. 203, 217 (2013).

⁹⁹ In later models, I represent the size of this effect by the formula $\tau w(Dc - Dr)$, that is the product of τ , the tax collected per unit of externality, times the number of units produced at equilibrium (widgets, w), times the difference in deadweight loss between the carbon tax and the revenue source it replaces. Note that w is not the number of units of carbon removed from the atmosphere, but instead the number of units produced. On the graph, that is the difference between quantity at equilibrium and quantity at the far right-hand side of the graph, where further reductions in carbon are no longer possible. One can think of this right-hand bound as being the first widget that rolls off the production line; W cannot buy fewer than zero widgets.

In this simple model, each widget produces one unit of carbon, so that w is equal to both the number of widgets and the number of units of carbon. More complex models can add an additional term for the number of units of carbon per widget.

For simplicity, I also assume throughout that the cost function of each producer is identical, and that the amount of revenues collected from the Pigouvian tax is small relative to the economy, such that the deadweight loss of each revenue source is effectively invaring. Finally, I assume a “closed” economy in which tax rates do not affect the flow of investments into or out of the country.

gets is \$100 per widget. W , a widget consumer, realizes a \$60 utility premium per widget absent the tax. Now suppose W is not that attentive to the tax; he perceives only 50% of the actual tax rate—or, according to my earlier terminology, has θ equal to 0.5. Since W thinks he will achieve a net premium of \$10 after tax, he buys the widget. That decision results in \$100 in revenue that the government would not have earned if the tax were fully salient.¹⁰⁰

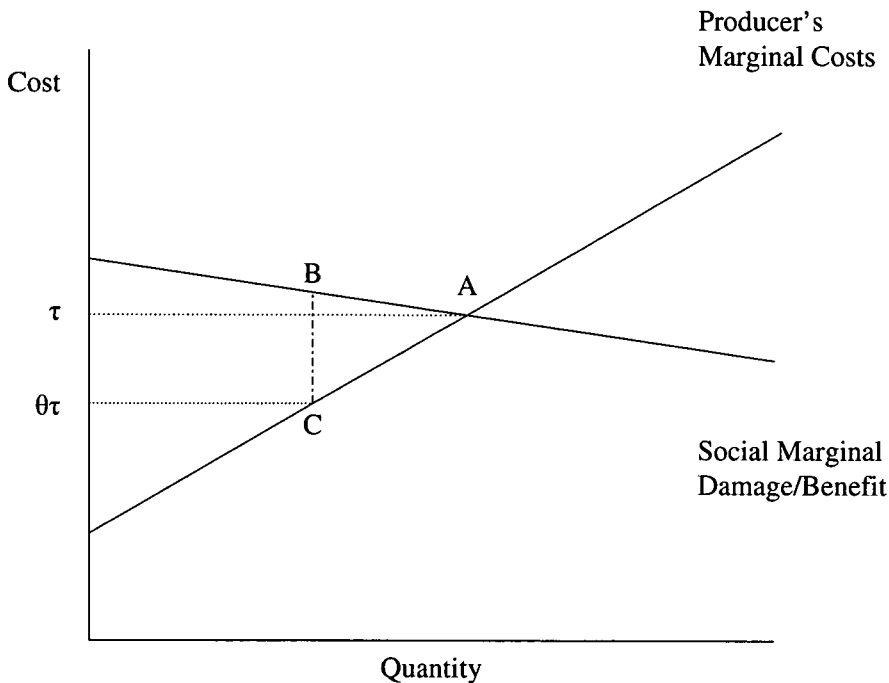
Though he notes that Pigouvian taxes should be more salient than those intended only to raise revenue, Goldin does not focus closely on the possibility that social damage from the resulting additional production of the externality could exceed any revenue gains.¹⁰¹ That is unlikely for the very first few units of overproduction. Consider the graph in Figure 2. The graph looks much like Figure 1, which was the graphical representation of the idea that the per-unit tax, τ , should be set at the price point where the cost of avoiding negative externalities and the social gains from avoiding them intersect, the “equilibrium.” But consumers do not respond fully to the tax; they act instead as if the tax were a lower price, $\theta\tau$. As a result, society produces too much carbon, with a net utility loss equal to area of the triangle, ABC .

Goldin’s revenue story is therefore most plausible for units very close to point A . Take *Marge* the marginal producer just to the left of A , whose widget surplus is \$99. If *Marge* thinks the tax is less than \$99, she sells a widget. Society loses the equivalent of \$1 in utility from that sale: *Marge* gains \$99, but her neighbors lose \$100. At the same time, *Marge* pays \$100 in taxes. Government reduces its other revenues by \$100. What are the savings from that? It depends on the deadweight loss of existing revenue sources. Estimates of the average deadweight loss of U.S. taxation run about 50% at their high end, though some research suggests the number may be rather lower or

¹⁰⁰ This example shows that the revenue impact in Goldin’s model results from increasing the tax base beyond w . W pays tax not on all widgets that could be built, but instead only those that are actually built—widgets where the price of mitigating the externality (which one can also think of as the manufacturer’s lost profit from not producing the widget, see Kaplow & Shavell, note 34, at 3 n.2) is greater than the perceived cost of the tax. In Figure 2, these are the widgets extending rightwards to the far-right bound from the intersection of $\theta\tau$ and the marginal cost line. As θ increases, this point shifts farther left, increasing the number of widgets produced and hence subject to tax. The number of widgets that are subject to tax as a result of the lower salience is the widgets sold between $\theta\tau$ and the old equilibrium point. One can therefore calculate the resulting revenue gain by computing the revenue benefits of taxing that number of new widgets. Using a standard quantity function, one can calculate that number readily enough, given the slope of that line, βc , and the knowledge that cost at the two end points of the line is τ and $\theta\tau$, respectively. Thus the sum will be $(\tau - \theta\tau)/\beta c$.

¹⁰¹ See Goldin, note 7, at 286-87.

FIGURE 2
PIGOUVIAN PRICING WITH PARTIAL SALIENCE



higher.¹⁰² It is too simple to say that society's utility gain is just 50% times \$100; the \$100 stick may also have created other distortions not seen in this picture, such as its potential effect on *Marge's* incentives to start a business and work hard at building it. But assuming that the difference between the deadweight loss of the stick and the tax it replaces is more than about 1% (on these assumptions), society has come out ahead. That is the "revenue effect" of low salience.

At some point the losses from excess production are going to exceed these revenue gains.¹⁰³ Take *Ginny*, *Marge's* cousin, who is close

¹⁰² See Charles L. Ballard, John B. Shoven & John Whalley, General Equilibrium Computations of the Marginal Welfare Costs of Taxes in the United States, 75 Am. Econ. Rev. 128, 128 (1985). For estimates implying higher and lower averages, see Martin Feldstein, Tax Avoidance and the Deadweight Loss of the Income Tax, 81 Rev. Econ. & Stat. 674, 678 (1999); Austan Goolsbee, Evidence on the High-Income Laffer Curve from Six Decades of Tax Reform, 2 Brookings Papers on Econ. Activity 1, 5-6, 13 (1999); Austan Goolsbee, Taxes, Organizational Form, and the Deadweight Loss of the Corporate Income Tax, 69 J. Pub. Econ. 143, 150 (1998). For discussion of Feldstein's methodology, see Raj Chetty, Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance, Am. Econ. J.: Econ. Pol'y, Aug. 2009, at 31 passim.

¹⁰³ In simple math terms, $U = \tau \left(\frac{\tau - \theta\tau}{\beta_c} \right) (D_c - D_r) - Du$ ($D_c - D_r$) - Du . That is, net utility from the low-salience tax (relative to the full-salience equilibrium) is equal to revenue gains less deadweight losses of underproduction, Du . But this equation is not very edifying. Since Du is just a triangle on the graph, one can derive Du as a function of basic

to point *C* on the figure. *Ginny* has a θ of 0.5, and she realizes a \$51 surplus per widget. She, too, pays the carbon tax, because she thinks it is only \$50. Extra carbon production from that decision costs society \$49 worth of utility: \$100 in damage less her \$51 profit. And it brings in \$100 in new revenues. But now, the deadweight loss split between the carbon tax and the corporate tax must be at least \$49 for this to be on net a social gain. Therefore it is unclear whether by itself low salience benefits society overall. Losses at the left side of the curve (close to point *C* in Figure 2) may or may not outweigh gains closer to equilibrium. The steeper the two lines, and the larger the θ discount, the more likely it is that losses will outweigh gains.

One other note to keep in mind is that my analysis omits complete consideration of the costs of consumer mistakes. Again, when taxes are hidden economic actors may pay more for goods than they expect, and this may mean that ultimately they end up with a different basket of purchases than they would have chosen under full information.¹⁰⁴ But as I mentioned before, this is in part an error of budgeting: That is, it depends to some extent on income effects.¹⁰⁵ Thus we will have

parameters, such as the slopes of the marginal cost (β_c) and social marginal damage curves (β_d), and of θ , through a bit of geometry and algebra (or, more generally, by taking the integral of the difference between the two lines and evaluating the result at our two end points). Ultimately, it is a slightly unsightly result,

$$U = \tau \left(\frac{\tau - \theta\tau}{\beta_c} \right) (D_c - D_r) - \left(\frac{\tau^2}{\beta_c} \left(\left(\frac{\beta_d}{\beta_c} \right) (\theta^2 - 2\theta + 1) + \theta^2 - 2\theta + 1 \right) \right)$$

This equation also confirms Goldin's result that the optimal salience can be less than full, assuming that there are some utility gains from revenue replacement. The first derivative

with respect to θ will be $-\frac{\tau^2}{\beta_c} (D_c - D_r) - \frac{\tau^2}{\beta_c} \left(\frac{\theta\beta_d}{\beta_c} - \frac{\beta_d}{\beta_c} + \theta - 1 \right)$. The second derivative is strictly negative, so setting the first derivative equal to zero to find the optimal θ yields:

$$\theta^* = \frac{\beta_d}{\beta_d + \beta_c} - \frac{\beta_c (D_c - D_r)}{(\beta_d + \beta_c)} + \frac{\beta_c}{\beta_d + \beta_c}$$

Assuming $D_c > D_r$, this implies that θ^* is between zero and one under many plausible parameters—for example, if the slope of the cost line is 2 and the slope of the damage line is 0.5, then $\theta = 0.6 - 0.2(D_c - D_r)$.

¹⁰⁴ See Chetty et al., note 60, at 1170-74.

¹⁰⁵ See id.; Goldin, note 7, at 276-77. Though Goldin acknowledges this point, his model appears to assume that the income and substitution effects of goods will be equally salient. If one holds household income constant, consumers may still lose through substitution effects, as when *Ginny* pays the carbon tax instead of buying a good without tax. See Goldin & Homonoff, note 62, at 307-10.

For modeling purposes, I allow for the impact of substitution effects on consumer welfare by treating the term D_c as being fixed at the amount of distortion that would be caused by the full price of the stick, rather than the discounted price perceived by the consumer. By assumption, deadweight loss is the amount a rational consumer and her trading partners lose when tax shifts the consumer to her next-preferred option. That same cost, then, is the largest loss that might be experienced by the consumer and her counter-parties if low salience leads her to choose the taxed good instead of the next-best option.

to defer full consideration of these mistakes to Section III.C.1 on income effects.

2. *Low Salience Sticks with Price Cross-Up*

Like Goldin, Gamage and Shanske also argue that low-salience taxes can be efficient, but they suggest that the government will need to increase taxes to achieve that result.¹⁰⁶ Again, revenue effects are the key. Gamage and Shanske suggest that price instrument rates can be adjusted—what I call “grossed up”—by, say, a factor of $1/\theta$, to offset any salience effects on the optimal level of negative externality produced.¹⁰⁷ If θ equals 0.5, then rates should be doubled. The offsetting shift is actually more efficient than full salience, they claim, because it multiplies the low-deadweight loss revenues derived from the stick.¹⁰⁸ For example, given marginal social damage of \$100, we would rather have a \$100 stick with taxpayers’ θ equal to 0.5 than a \$50 stick with θ equal to 1. Both result in the exact same amount of externality produced. But the higher-priced stick brings in twice as much in low-distortion revenues (assuming away income effects). Therefore, unlike the scenario in which the stick’s rate is not adjusted, there is no deadweight loss from excess production of the negative externality.¹⁰⁹

Gamage and Shanske acknowledge but do not fully explore some important limits on their result. For one, as they note, the salience of a tax very likely varies among taxpayers.¹¹⁰ In some settings this heterogeneity implies that offsetting adjustments are not optimal.¹¹¹ For example, suppose there are two groups of taxpayers, “sophisticates” with θ equal to 1 and “naïves” with θ equal to 0.5. If both face the same stick price, and policymakers increase that price to fully deter naïves, the sophisticates will be overdeterred. Consider Figure 3.

¹⁰⁶ See Gamage & Shanske, note 7, at 71-74.

¹⁰⁷ Id. at 72-73; see also Allcott et al., note 7, at 85 (“[W]hen the average marginal consumer undervalues energy efficiency,” increasing an energy tax yields “higher welfare gains . . . than would be expected in the externality-only case.”).

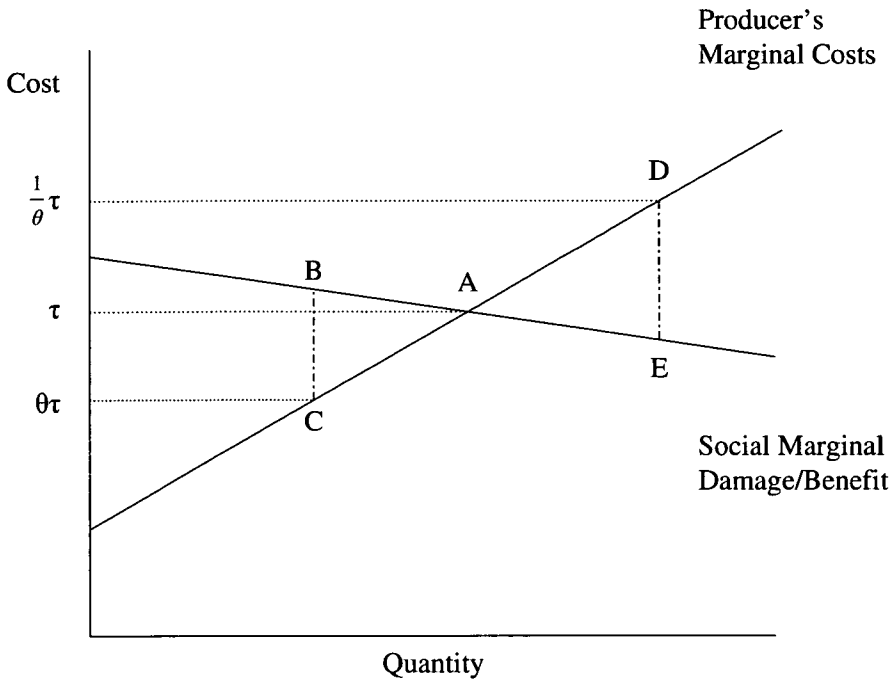
¹⁰⁸ See Gamage & Shanske, note 7, at 72-73.

¹⁰⁹ See id. at 72-73.

¹¹⁰ See id. at 77.

¹¹¹ See Allcott et al., note 7, at 87 (analyzing implications of heterogeneous undervaluation for optimal energy tax policy).

FIGURE 3
 PIGOUVIAN TAX, PARTIAL SALIENCE, FULL GROSS-UP,
 HETEROGENEOUS PRODUCERS



In this diagram, the government is imposing a carbon tax that is grossed up to fully offset the lowest salience in the population. In other words, even *Ginny*, with θ equal to 0.5, will perceive the tax as costing her \$100. That means the real tax is $1/\theta$ times \$100, or \$200 per widget. Suppose we have another widget consumer, *Maxwell*, who is sophisticated enough to recognize the real price. Even if *Maxwell* could earn \$199 in profit by purchasing a widget, he will not, because that would also cost him \$200. When *Maxwell* decides to halt widget consumption, society loses. It does gain \$100 from reduced carbon emissions. But it also loses *Maxwell's* \$199 profit. So there is a net \$99 deadweight loss. The sum of these losses, from *Maxwell* and his smart friends, is the triangle *CDE*.¹¹²

¹¹² One can represent the offsetting revenue effects and overproduction losses with an equation similar to the equation earlier. Here, the increased tax will be paid on every widget sold by the producer (that is, on every unit of carbon produced). So set gains from new revenues as $\frac{1}{\theta} \pi w (D_c - D_r)$, where w is simply the number of widgets sold at equilibrium. But this tax is only paid by naïve taxpayers. So weight it by P_n , the naïve population. The losses from overdetering emissions will similarly be weighted by the sophisticated population, P_s . Using the same geometry techniques employed to determine the area of the deadweight loss triangle for undermitigation, one can find the area of the deadweight loss triangle of overmitigation. Putting these together yields:

Gross-up also costs the government some revenue in a world with sophisticated taxpayers. Consider *Maxwell* again. If the carbon tax were \$100, and *Maxwell's* costs of mitigating carbon in his widgets were above \$100, *Maxwell* would emit the carbon and pay the tax. But if the tax is \$200, and his cost of mitigating is, say, \$199, he will eliminate the carbon and avoid the tax. Thus the government loses revenues, relative to a full-salience world, on the widgets for which the cost of mitigating falls between τ and $\frac{1}{\theta}\tau$.¹¹³ By raising the tax rate, government also increases the distortiveness of the stick for sophisticates who perceive the higher grossed-up price, diminishing the dead-weight-loss savings of the resulting stick revenues.¹¹⁴

It appears that in many instances low salience with gross-up produces greater social welfare than a fully salient stick.¹¹⁵ For instance, if there are many naïve taxpayers for each sophisticate, the revenue effects of the gross-up could significantly outweigh the loss from any overproduction. It is less clear whether grossing up a low-salience stick is superior to simply allowing the salience to remain low. While there are revenue losses among sophisticates from gross-up, the revenue effect from raising the tax paid by naïves can be several multiples of the revenue effect from low salience alone.

The relative sizes of the losses from missing the optimal amount of pollution are also important inputs in evaluating the comparative merits of the two approaches. In Goldin's model, producers emit too much carbon,¹¹⁶ while in Gamage and Shanske's plan, they mitigate too much, producing a less-than-optimal amount of pollution.¹¹⁷ Stan-

$$P_n \left(\frac{1}{\theta} \tau w(D_c - D_r) \right) - P_s \left(\frac{\tau^2}{2\beta_c} \left(\frac{\beta_d}{\beta_c} \left(\frac{1}{\theta^2} - \frac{2}{\theta} + 1 \right) + \frac{1}{\theta^2} - \frac{2}{\theta} + 1 \right) \right)$$

It is worth noting that the size of the deadweight loss triangle here is bounded on the right by the maximum possible reduction. At some point, producers have eliminated all externalities, or taxes have risen so high that no widgets can be sold profitably, and producers go bankrupt.

¹¹³ One can add this revenue loss to the earlier equation. As before, one can calculate the quantity of widgets that fall in between two points using the standard quantity formula, the slope of the cost curve, and the knowledge that the cost at the two end points must be τ and $\frac{1}{\theta}\tau$. That calculation yields a quantity equal to $(\frac{\tau}{\theta} - \tau)/\beta_c$. To find the total utility cost of the lost revenue, one then would weight that quantity by the sophisticate population and multiply by the utility cost of forgone replacement revenue, $P_s \left((\frac{\tau}{\theta} - \tau)/\beta_c \right) (D_c - D_r)\tau$.

¹¹⁴ Gross up also increases the welfare loss for consumers who make poor choices. As before, the upper bound of that welfare loss can be captured by the deadweight loss for the full tax price, D_c .

¹¹⁵ Adding together with the earlier formulae:

$$U = P_n \left(\frac{1}{\theta} \tau w(D_c - D_r) \right) - P_s \left(\frac{\tau^2}{2\beta_c} \left(\frac{\beta_d}{\beta_c} \left(\frac{1}{\theta^2} - \frac{2}{\theta} + 1 \right) + \frac{1}{\theta^2} - \frac{2}{\theta} + 1 \right) \right) \\ - P_s \left(\left(\frac{\tau}{\theta} - \tau \right) / \beta_c \right) (D_c - D_r)\tau$$

¹¹⁶ See Goldin, note 7, at 286.

¹¹⁷ See Gamage & Shanske, note 7, at 73.

dard economic theory suggests that this latter case, overmitigation, typically produces bigger deadweight losses.¹¹⁸ According to that theory, price instruments are most preferable to other tools when the marginal social damage curve is fairly flat: Each additional unit of the public bad is not much more harmful than the one before.¹¹⁹ That suggests that when we are using price instruments, overdeterrence is usually the more serious concern.¹²⁰ It is worth keeping in mind, though, that with gross-up only sophisticates overmitigate, while with no gross-up there is excess pollution from all naïves.

In short, there are good reasons to think that in some situations full gross-up is a better option. But, as I now show, these are not the only two choices.

3. *Low Salience Sticks with Partial Gross-Up*

Although fully grossing up stick prices to account for low salience is intuitive, government can choose many other pricing strategies. Indeed, as Gamage and Shanske acknowledge, in some cases political opposition to a stick will constrain its price below the value that would fully incentivize naïve producers.¹²¹ Government can set the price of

¹¹⁸ See, e.g., Hepburn, note 11, at 231-32.

¹¹⁹ The typical reason offered in the literature for preferring price instruments is that we think the risk of excess cost is greater than the risk of excess externalities. See, e.g., Helfand et al., note 28, at 279; cf. Stavins, note 34, at 889 (concluding that, “if marginal abatement costs are flat relative to marginal benefits, then a quantity instrument is more efficient than a price instrument”). When the marginal social damage curve is flat, we prefer to avoid overdeterrence, since the cost of avoiding each additional unit typically rises, but the benefits of avoiding that unit are not changing much. Therefore, if observing that officials have chosen a price instrument, one might reasonably conclude that they believe the marginal social damage curve is fairly flat. Kaplow & Shavell, note 34, at 7-10, argue that taxes can account for this possibility, if the tax rate is allowed to vary across taxpayers or over time. But Ben-near and Stavins argue that firm responses to government policy occur over long periods, such that frequent policy changes may be unrealistic. Ben-near & Stavins, note 50, at 122. They suggest in that situation that a combination of price and quantity instruments might be optimal regardless of the shape of the marginal social damage curve. See *id.*

Another argument for price instruments arises if the combination of quantity regulation plus tradable permits is not possible. For example, pollutants that are more harmful when concentrated in one place, such as mercury, are usually not good candidates for trading regimes. In those instances the mere fact that officials have chosen a price instrument regime would not tell us much about the slope of the marginal social damage curve.

¹²⁰ See Sterner, note 4, at 212. Again, Kaplow & Shavell, note 34, at 7-10, argue that the costs of overdeterrence could be mitigated with flexible tax rate schedules. If that were the case, could impose lower taxes on sophisticates to avoid some of the costs of grossing up. But this would require the government to be able to identify sophisticated taxpayers. As Kaplow and Shavell implicitly acknowledge, when rates vary according to observable characteristics of the taxpayer, there are often opportunities for strategic behavior, which will tend to diminish the efficacy of any flexible schedule. *Id.* at 6 n.7.

¹²¹ See Gamage & Shanske, note 7, at 73.

a stick at any multiple of θ , though the most logical values are probably partial gross-ups, such that the price falls somewhere between one and $1/\theta$ times the marginal social damage. So, for instance, if naïve producers have θ equal to 0.5, and marginal social damage is \$100, the government might set its stick price at \$150, rather than the \$200 per unit that would cause naïves to fully internalize their carbon damage. I argue that, rather than a bug, this result actually would be a feature: Partial gross-up is likely to be more efficient than other responses to low salience.

Though the math here is a bit messy, the intuition in favor of partial gross-up is straightforward and familiar to most readers acquainted with tax policy. If government does not set the stick price high enough to cause a naïve producer to fully internalize her externalities, she will overproduce them.¹²² At the same time, by lowering the tax price facing sophisticates, government reduces the extent to which they work too hard to reduce externalities.¹²³

In other words, partial gross-up reduces one distortion at the cost of increasing another. Why would that be helpful? The deadweight loss of each distortion bears an exponential relationship to the size of the distortion.¹²⁴ That means that the total loss from two small distortions of size X will be much less than the lost utility caused by a single distortion of size $2X$. In the case of the carbon tax example, reducing gross-up allows one to shave off the biggest piece of the overdeterrence deadweight loss triangle in exchange for a much smaller new triangle, as illustrated in Figure 4.

¹²² Recall the earlier equation in which the deadweight loss of underdeterrence was given as $\frac{\tau^2}{2\beta_c} \left(\left(\frac{\beta_d}{\beta_c} \right) (\theta^2 - 2\theta + 1) + \theta^2 - 2\theta + 1 \right) + \theta^2 - 2\theta + 1$. Now let x represent some adjustment in the degree of gross-up, such that the tax imposed is $\frac{1-x}{\theta} \tau$. The area of the resulting underdeterrence deadweight loss triangle can then be

given as:

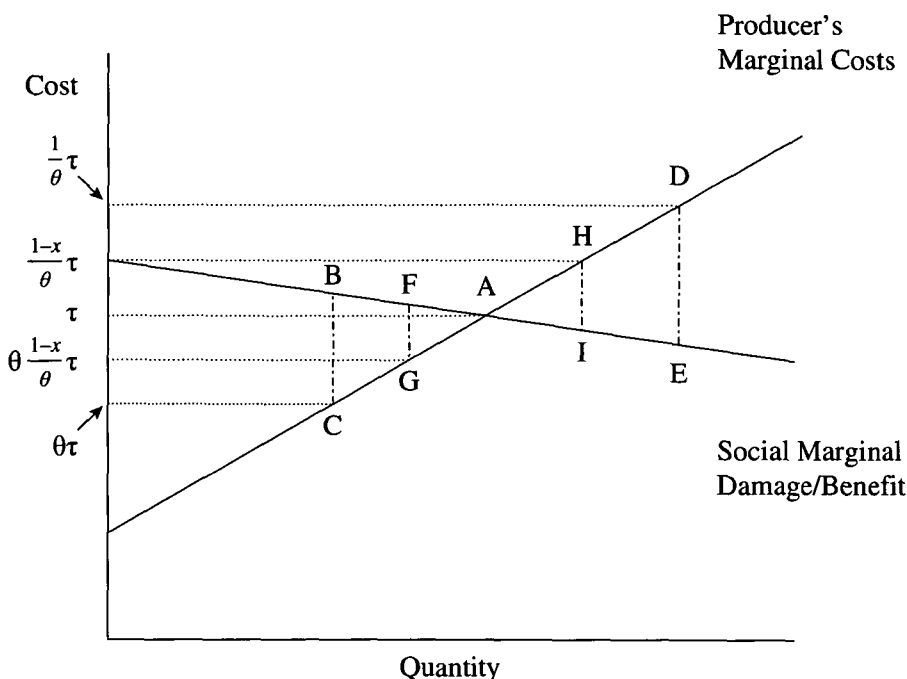
$$\frac{x\tau}{2\beta_c} \left(\frac{\beta_d}{\beta_c} x\tau + x\tau \right)$$

¹²³ Again, using simple geometry and algebra, one can calculate the resulting deadweight loss. If x is the adjustment in the degree of gross-up, such that the tax imposed is

$\frac{1-x}{\theta} \tau$, then deadweight loss of overdeterrence is: $\frac{x\tau}{2\beta_c} \left(\frac{\beta_d}{\beta_c} x\tau + x\tau \right)$

¹²⁴ Chetty et al., note 60, at 1173.

FIGURE 4
 PIGOUVIAN TAX, PARTIAL SALIENCE, PARTIAL GROSS-UP,
 HETEROGENEOUS PRODUCERS



Admittedly, there is considerably more complexity in the problem than is captured in this simple intuition. Most importantly, by reducing gross-up, we also potentially diminish the revenue collected by the tax. There are two offsetting revenue effects. For one, we are reducing the amount of tax collected on each unit of carbon. But, at the same time, since naïve producers are no longer fully deterred, they will now produce more taxable units of carbon than at the full gross-up equilibrium.¹²⁵ For example, if *Ginny*, with θ equal to 0.5, faced a tax of $1/\theta$ times the marginal social damage of \$100, she would produce no widgets where the costs of eliminating carbon were more than \$100. But if the price is only \$150 per unit, she will perceive the per-unit tax as \$75. She then will produce widgets for which it would cost her between \$75 and \$100 to eliminate the carbon, and she will pay tax of \$150 on each of those widgets. Similarly, sophisticated producers will also produce more widgets because the tax cost of doing so has

¹²⁵ Under full gross-up naïve producers produce widgets at the full-salience equilibrium amount, so this change in quantity is also the change relative to full salience. Relative to either equilibrium, then, partial gross-up to a tax of $1 - x/\theta$ increases the number of units produced by $P_n\left(\frac{x\tau}{\beta_c}\right)$, with welfare effects $P_n\left(\frac{x\tau}{\beta_c}\right)\left(\frac{1-x}{\theta}\tau\right)(D_c - D_r)$.

declined.¹²⁶ In other words, partial gross-up lowers rates but expands the tax base relative to full gross-up.¹²⁷

Overall, intuition suggests that the beneficial effects of reducing total externality-related costs—relative to the costs under full or no gross-up—should dominate any sacrifice from diminished revenues.¹²⁸

While both quantities may depend on exponential functions of the tax rate, lost revenues from declining rates are partly offset by revenues from an expanding base. Lower rates also reduce the distortiveness of the stick, improving its per-dollar deadweight loss gains over a corporate tax. Thus, where externality-related costs are large, reducing the stick price from full to partial gross-up should tend to improve welfare. If there are many more naïve producers than sophisticates this gain will be muted, however. As government continues to reduce the tax rate, the benefits from cutting externality costs further become small relative to the possibility of revenue gains, and so there is plausibly an optimal price falling somewhere between marginal social damage and $1/\theta$ times that amount.¹²⁹

¹²⁶ Among sophisticates, the change in quantity from gross-up to partial gross-up is $P_s \left(\frac{x\tau}{\theta} / \beta_c \right)$.

¹²⁷ Overall, as x increases the total change in revenue relative to full gross-up can be given as the sum of the contending revenue effects:

$$P_n \left(\frac{x\tau}{\beta_c} \right) \left(\frac{1-x}{\theta} \tau \right) + P_s \left(\frac{x\tau}{\theta} / \beta_c \right) \left(\frac{1-x}{\theta} \tau \right) - P \left(\frac{x}{\theta} \tau w \right).$$

¹²⁸ Building on earlier equations, and including an additional term for underdeterrence, the net utility of a partially grossed-up tax, relative to a full-salience equilibrium, can be given as:

$$\begin{aligned} U = & P \left(\frac{1-x}{\theta} \tau w (D_c - D_r) \right) + P_n \left(\frac{x\tau}{\beta_c} \right) \left(\frac{1-x}{\theta} \tau \right) (D_c - D_r) \\ & - P_s \left(\frac{\tau^2}{2\beta_c} \left(\left(\frac{\beta_d}{\beta_c} \right) \left(\frac{x^2 - 2x + 1}{\theta^2} - \frac{2 - 2x}{\theta} + 1 \right) + \frac{x^2 - 2x + 1}{\theta^2} - \frac{2 - 2x}{\theta} + 1 \right) \right) \\ & - P_s \left(\left(\frac{\tau - x\tau - \theta\tau}{\theta} \right) / \beta_c \right) (D_c - D_r) \tau - P_n \left(\frac{x\tau}{2\beta_c} \left(\frac{\beta_d}{\beta_c} x\tau + x\tau \right) \right). \end{aligned}$$

In this equation, the first two terms represent, respectively, the revenue effects of raising the tax rate and broadening the base. The next term captures the losses from overdetering sophisticates, and the fourth term captures revenues lost from sophisticates who mitigate rather than paying the higher tax. The final term captures residual deadweight loss from naïve producers who are not fully deterred by the partially salient stick.

Note that because the two curves are bounded to the left and right by full and zero emissions, respectively, it is also possible that there will be corner solutions in which the expected deadweight loss is truncated. In these situations partial gross-up may no longer be optimal.

¹²⁹ To test this mathematically, one can treat θ as a parameter and attempt to find an optimal x (reduction in gross-up ratio) given any θ . By taking $\frac{\partial U}{\partial x}$ in the equation in note 128 (and noting that the second derivative is strictly negative), setting that result equal to zero, and solving for x , one can, after a great deal of algebra, conclude that:

A very similar approach, and one that is not mutually exclusive with partial gross-up, would be for government to combine gross-up with attempts to change the salience of its stick. If government succeeds in making the stick more salient, it will diminish its need to gross up the price of the stick. That means it will collect less revenue from naïve producers. At the same time, it will distort the incentives of sophisticated producers by a smaller amount than it would if $1/\theta$ were larger. In short, the math of adjusting salience looks quite similar to partial gross-ups. This equivalence allows the government to choose the technique or combination of techniques that are best achievable with current salience-altering technologies and within existing political constraints on tax rates.

Partial gross-up or adjusted salience also is flexible enough to adapt to even greater heterogeneity in the population than I have considered so far. Realistically, the population will not be neatly divided into naïve and sophisticated producers, but instead will likely include individuals with a range of θ values. My analysis here suggests that it will usually not be ideal for government to set tax rates so as to optimize the response of the *most* naïve individuals, but instead to allow some underproduction by those producers.

Despite these caveats, I agree with the central findings of Gamage & Shanske and Goldin that low salience can be efficient even for Pigouvian taxes. Further adjustments to the design or rate of the stick, however, may achieve even more efficient outcomes than they suggest. As shown below, including income effects in the mix may also add some important wrinkles.

B. Extension I: Charitable Contributions and Other Positive Externalities

So far I have focused, as price-instrument theorists traditionally have, on the use of Pigouvian pricing to control negative externalities. How much of that discussion would also apply to positive externalities? For instance, consider credit-hours of education. These credit-hours benefit not only the purchaser, but also future employers (think of the law firm that need not train its first-year associates to do the

$$x > 0 \text{ iff } \frac{x^2}{\beta_c} > \tau w,$$

where w again is the number of units of the polluting good produced at equilibrium. That

quantity can be restated as $w = \frac{\tau^2}{\beta_c} - \frac{\tau\alpha}{\beta_c} - \tau Q_{\max}$, where α is the y-intercept of the cost curve, and Q_{\max} is the quantity produced with no reduction in emissions (in essence, the intercept of the cost curve and the right-hand side of the diagram). Therefore, assuming that α is greater than or equal to zero (and that τ , which is a per-unit quantity rather than a tax rate, is greater than 1), x is strictly positive. Q.E.D.

tasks law school has prepared graduates to do) and society as a whole. What if we imposed a penalty on young people who skip college?¹³⁰ Or, less hypothetically, what if there were a fee imposed on individuals above a certain income threshold who fail to obtain health insurance?¹³¹

1. *Positive Externalities and Sticks*

Flipping the externalities government targets from negative to positive should not change the low-salience analysis I have done so far. In one sense, semantics are all that separate positive externalities from negative. We can describe the actions we want our widget producers to undertake either as cutting their negative-externality production, or instead as increasing a new positive externality, pollution reduction. Either way, the marginal social gains from an additional unit of pollution taken out of the atmosphere would be the same. The same for education: With each hour of education bought or not, a student contributes either to a well-educated population (positive externality) or to an ignorant one (negative).

As I argued earlier, though, in their archetypical cases positive and negative externalities differ in their relation to the status quo.¹³² To produce more positive externalities, society usually has to build more “stuff”: hospitals, research labs. To reduce negative externalities, it must diminish something, such as a polluting industrial process, or traffic, or cigarettes consumed. This divergence between “more stuff” and “less stuff” informs my discussion of income effects. It also affects government’s ability to solve private market failures directly.

As Emmanuel Saez explains, government can respond to private overproduction of some public good by reducing its own output.¹³³ His example is charitable contributions.¹³⁴ If donors give too much to charity, government can either reduce its subsidy for charity, or instead simply produce fewer of the goods that government produces in common with the charitable sector. If there are too many private

¹³⁰ If this seems politically infeasible, consider that education deferments from the draft during the Vietnam War were similar in structure to the fee I just mentioned: They imposed higher costs, albeit non-monetary costs, on a cohort of young people who chose not to attend college. Whether the fact that the costs were nonmonetary should change the normative analysis of them is a complicated question I leave for elsewhere. See Brian Galle, *Tax, Command. . . Or Nudge? Evaluating the New Regulation*, 92 *Tex. L. Rev.* 837 (2014).

¹³¹ See IRC § 5000A(b)(1).

¹³² See text accompanying notes 56-57.

¹³³ See Emmanuel Saez, *The Optimal Treatment of Tax Expenditures*, 88 *J. Pub. Econ.* 2657, 2659 (2004).

¹³⁴ *Id.* at 2660.

schools, government could reduce investment in public schools. Thus, Saez argues, government can use its own output to adjust for imperfectly-targeted subsidies for charity.¹³⁵ In some cases this second instrument will allow government to obtain better results than using the price instrument alone.

Saez's analysis implies that it is easier for government to correct overproduction of the archetypical positive externality than it is to do the same for overdeterrence of the classic negative externality. Government does not produce lower cigarette consumption. If sticks aimed at curtailing smoking are priced too high, there is no obvious government output that can be lowered in response.¹³⁶

The upshot of this analysis is that the carrot and stick cases may not be perfectly symmetrical. Grossing up sticks typically results in over-incentivizing sophisticates. If government has other mechanisms available to offset this effect, its impact can be mitigated, potentially increasing the appeal of Gamage & Shanske's proposal. Saez's theory suggests that possibility is more plausible for most positive externalities.

2. *Positive Externalities and Inefficient Carrots*

In any event, it is unusual in our society to use sticks to encourage positive externalities. Politics and law contribute to that outcome, as I have explained elsewhere.¹³⁷ It is also the case that sometimes carrots are more efficient, usually because of income effects. If the socially desirable outcome requires the producer to buy something—which we typically observe in the positive-externality case—we often want her to have more money, so that she will want and be able to afford more of it. As I explained previously,¹³⁸ that is likely the best argument for why subsidizing donations to schools is a better way to educate the public than penalizing people who skip class or refuse to donate.¹³⁹ What then are the effects of low salience on, say, the charitable contribution deduction?

¹³⁵ Id. at 2667.

¹³⁶ To the extent that government does currently subsidize some tobacco production, it could reduce that subsidy, but in my analysis reducing a subsidy is the equivalent of a stick. In addition, governments may also use nonprice tools such as indoor smoking bans to reduce tobacco consumption. If these exist alongside the price instrument, then potentially government has flexibility similar to that which it has when directly producing the good.

¹³⁷ E.g., Galle, note 4, at 840-45.

¹³⁸ See Section II.B.

¹³⁹ See Galle, note 4, at 832-38.

Perhaps counter-intuitively, in many settings low-salience carrots could be superior to their more visible counterparts.¹⁴⁰ One reason, as for sticks, is that lower salience potentially improves the revenue effects of the price instrument: If individuals perceive the reward of the carrot as too low to motivate them, the government does not need to pay out. As I mentioned earlier,¹⁴¹ carrots also carry other welfare costs, as well, including the likelihood that they will encourage future bad acts, or crowd out future good acts, by producers.¹⁴² To be sure, diminishing the salience of the carrot may cause over- or under-production of externalities in the short run. But especially in the case of negative externalities, it is quite possible that the optimal salience of the carrot is zero. That is, in some situations it would be better for society if the carrot had no effect than for society to continue to pay for the effects the carrot currently produces.

At first glance this is a puzzle: Why would government want an instrument it enacted to be invisible to the public? One answer, as I explore below¹⁴³, is that the instrument might only be hidden from some actors, and this partial salience could be more effective than full salience. But even where that is not the case, salience could potentially be a second-best solution to political failures. Carrots, again, are often the product of defective politics and unfortunate legal doctrine. Because carrots usually embody a transfer of wealth from all taxpayers to a small, relatively cohesive group (the externality producers), lobbying and policy decentralization both tend to favor carrots.

Alternately, the carrot might have been enacted at a time when the consequences of that decision were not fully understood. With the passage of time, society regrets its decision, but the subsidy now is too politically entrenched for repeal. Lowered salience at least would diminish the ongoing cost of the locked-in error. The home mortgage interest deduction provides one possible example.

Salience could allow an actor who is less subject to political pressures to partially remedy the bad decisions of another. Many actors can influence the impact of a price instrument, including different branches and tiers of government. Thus federal policy might—per-

¹⁴⁰ See Lily L. Batchelder, Fred T. Goldberg, Jr. & Peter R. Orszag, *Efficiency and Tax Incentives: The Case for Refundable Tax Credits*, 59 *Stan. L. Rev.* 23, 48-49 (2006) (arguing that inefficient subsidies should be targeted to actors whose behavior is the least elastic); cf. Martin Feldstein, *Tax Avoidance and the Deadweight Loss of the Income Tax*, 81 *Rev. Econ. & Stat.* 674, 674-77 (1999) (arguing that tax avoidance may improve welfare if the income tax system has many inefficient subsidies). But cf. Stavins, note 34, at 889 (noting that eliminating subsidies for inefficient behavior may be the best current approach to pollution abatement).

¹⁴¹ See Section II.B.

¹⁴² Galle, note 4, at 833-36.

¹⁴³ See Section III.E.

haps in combination with private third parties—use salience to soften the costs of carrots created by states, or the executive branch could diminish the appeal of congressional giveaways. I do not mean to suggest that this combination of events is necessarily intentional; public choice factors and low salience may accidentally combine to mitigate the damage from carrot overgrowth. Federal efforts to encourage homeowners to refinance their underwater homes appear to be, for good or ill, a recent example. Banks have apparently tried to make the program as opaque as possible, and relevant federal regulators do not seem to have pushed for greater transparency.¹⁴⁴

3. *Positive Externalities and Efficient Carrots*

Not all carrots would be better off hidden under the rug. There are some decent arguments that the charitable contribution deduction, for example, is a better choice than penalizing the selfish.¹⁴⁵ But there is also some evidence that donors are not fully aware of the benefits of the deduction.¹⁴⁶ Should policy makers “gross up” the price of a low-salience carrot to maintain its substitution effects?

Even in the case where the optimal salience of the carrot is greater than zero (that is, the carrot is actually good policy), the answer is still “probably not.” Grossing up of course magnifies the revenue costs of the carrot. To the extent that some producers are more aware of the carrot than average, an undifferentiated gross-up would wastefully over-incentivize them, perhaps including their incentives to play games with future policy. Further, gross-up multiplies the wastefulness of subsidizing inframarginal producers. Gross-up is welfare-improving only to the extent that the benefits of incentivizing *marginal* producers exceeds its combined costs.¹⁴⁷

Again, there is a good chance that partial gross-up would do better. Though naïve producers will not produce at optimal levels, the waste from over-production by sophisticates is much reduced. The revenue

¹⁴⁴ See, e.g., Jean Braucher, *Humpty Dumpty and the Foreclosure Crisis: Lessons from the Lackluster First Year of the Home Affordable Modification Program (HAMP)*, 52 *Ariz. L. Rev.* 727, 776-78 (2010).

¹⁴⁵ See Galle, note 4, at 839-40.

¹⁴⁶ Goldin & Listokin, note 70, at 7-8; see Jon Bakija & Bradley T. Heim, *How Does Charitable Giving Respond to Incentives and Income? New Estimates from Panel Data*, 64 *Nat'l Tax J.* 615, 636 (2011) (suggesting that their results could be explained by variations in “how salient the price changes” of charitable giving are).

¹⁴⁷ Putting this idea in math is a bit misleading, because it implies somewhat more precision than is actually possible. Several of the important factors in increasing the value of carrots, such as the waste of funds devoted to inframarginal producers, cannot easily be modeled in this framework. But as a rough approximation, consider (where τ again represents the amount of the Pigouvian price, which here is the subsidy, and β is the slope of the marginal social benefit curve):

cost of the subsidy for each unit produced also falls. Once more, deadweight losses increase exponentially and revenue effects are mixed. If the naïve population is large, intuition suggests that shrinking the losses from underproduction would, to a point, often be worth any revenue cost.

C. *Income and Output Effects*

Another extension of my basic results would be to factor in income effects. Until this point I have assumed away any income effects of sticks or carrots. I have also ignored the possibility that either instrument might similarly affect the outputs of firms. I remedy that oversight here.

1. *Individuals and Income Effects*

In many cases, but not all, the income effects of a single price instrument will be small relative to its substitution effects. Prior commentators therefore have largely assumed away income effects, or presumed that they will be irrelevant if the government bundles them with offsetting taxes or lump-sum payouts.¹⁴⁸ But in some settings income effects can be quite substantial. For example, analysts estimated that legislation resembling the carbon-tax bill passed in 2009 by the U.S. House of Representatives¹⁴⁹ would have consumed more than 5% of

$$U = P_n \left(\frac{\tau^2}{2\beta_c} \left(\left(\frac{\beta_b}{\beta_c} \right) (\theta^2 - 2\theta + 1) + \theta^2 - 2\theta + 1 \right) \right) \\ - P_s \left(\frac{\tau^2}{2\beta_c} \left(\left(\frac{\beta_b}{\beta_c} \right) \left(\frac{1}{\theta^2} - \frac{2}{\theta} + 1 \right) + \frac{1}{\theta^2} - \frac{2}{\theta} + 1 \right) \right) - P \left(\frac{1}{\theta} \tau w D_r \right) - P_n \left(\frac{D_r}{\beta_c} (\tau^2 - \theta \tau^2) \right).$$

What this equation tells us is that the net welfare effect of grossing up carrots to account for diminished salience reflects the gains from eliminating under-production by naïve producers, reduced by the costs of over-incentivizing sophisticates, and also reduced by the greater costs of subsidies. Subsidy costs increase for each widget produced by the entire producer population (the penultimate term on the right-hand side), and also increase with the number of new widgets produced by naïve producers (the rightmost term). By assumption, the first term on the right-hand side is greater than the last (that is, the carrots are worth their cost under full salience), but of course it does not follow that the first term exceeds all three costs in combination. I consider the impact of income effects in more detail in Section III.C.

¹⁴⁸ See Helfand et al., note 11, at 280; Parry & Oates, note 96, at 604.

¹⁴⁹ Actually, the bill established a system of tradable permits whose economic effect was similar to a carbon tax. American Clean Energy and Security Act of 2009, H.R. 2454, 111th Cong.; Cong. Budget Office, Cost Estimate: H.R. 2454 American Clean Energy and Security Act of 2009, at 20 (2009). On the economic similarity of cap and trade to a carbon tax, see Helfand et al., note 28, at 281.

household income for some poor families.¹⁵⁰ Data also suggest that very wealthy households frequently hit a 50%-of-income cap on the deductibility of their charitable contributions, implying that the dollar value of the deduction for those taxpayers is about 20% of their income.¹⁵¹ Household budget changes of this magnitude warrant some consideration.

As I showed earlier, income effects do not change the optimal price of a stick or carrot but do change the amount of externality that is produced.¹⁵² Carrots shift the typical producers' cost curve to the right, assuming we are dealing with normal goods. Sticks on normal goods in effect would shift the cost curve leftwards, as producers would demand fewer units of the good for any given cost. Salience can interact with these shifts in at least two distinct ways.

First, salience could diminish income effects together with substitution effects. If households are unaware of the value of their charitable contribution deduction, they may also miscalculate their after-tax budget for the year, and therefore assume that their bank balance will be smaller than their tax refund will show. Yet we often want charitable households to feel wealthier so they will donate more.¹⁵³ Or, in the other direction, households that are unaware of the bite a carbon tax takes out of their budgets may consume more than they would if the income effect were more salient.¹⁵⁴ In both these cases, salience is dampening a desirable shift in the cost curve. The opposite is also possible, such as would be the case if we used sticks to encourage charity.

Second, if income effects are relatively more salient than are substitution effects, they might affect the efficiency analyses discussed

¹⁵⁰ See Cong. Budget Office, *Trade-Offs in Allocating Allowances for CO2 Emissions 2* (2007); Terry Dinan & Diane Lim Rogers, *Distributional Effects of Carbon Allowance Trading: How Government Decisions Determine Winners and Losers*, 55 *Nat'l Tax J.* 199, 212 tbl.4 (2002).

¹⁵¹ Miranda Perry Fleischer, *Generous to a Fault? Fair Shares and Charitable Giving*, 93 *Minn. L. Rev.* 165, 172-76 (2008). Admittedly, households that can smooth their income over time likely behave as though their budget constraint is equal to lifetime rather than annual income. Therefore the 50% cap may be a rather smaller percentage of wealthy donors' effective budgets. But many donors hit the cap year after year, *id.*, which suggests that the income effect of donations is still substantial.

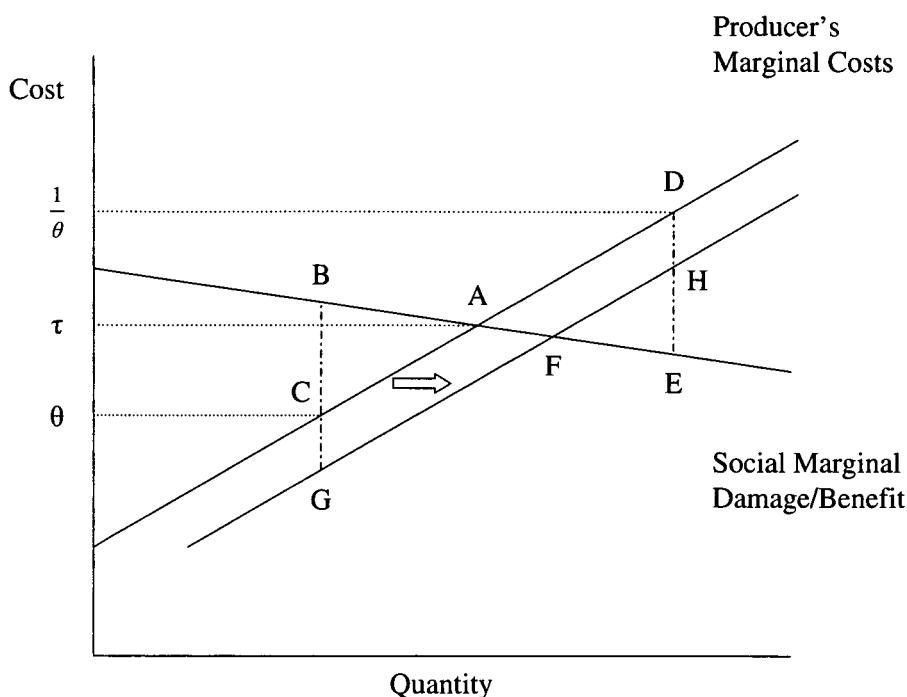
¹⁵² See Figure 1 and text accompanying notes 49-53.

¹⁵³ Once more, government officials typically cannot set subsidy amounts to take accurate account of the income effects of their carrot. At best, they can try to make sure that income and substitution effects are pointing in the same direction. In that situation, although there is some risk of overproduction, the officials at least can be assured that the income effects will not totally cancel out the substitution effect of the carrot.

¹⁵⁴ Cf. Brian Galle & Manuel Utset, *Is Cap-and-Trade Fair to the Poor? Shortsighted Households and the Timing of Consumption Taxes*, 79 *Geo. Wash. Univ. L. Rev.* 33, 57-60 (2010) (examining effect of energy-tax salience on energy consumption).

above.¹⁵⁵ In essence, by shifting the cost curve, income effects can affect the deadweight loss of under- or over-incentivizing externalities, as Figure 5 illustrates.

FIGURE 5
PIGOUVIAN TAX, FULL GROSS-UP, HETEROGENEOUS
PRODUCERS, WITH INCOME EFFECTS



This figure illustrates a partially hidden carrot—say, the charitable-contribution deduction—with fully-salient income effects.¹⁵⁶ As the cost curve shifts rightwards, the deadweight loss from naïve taxpayers, represented by triangle *ABC*, expands to the triangle *FBG*. The losses that would be produced by grossing up the tax rate to $1/\theta$, as represented by triangle *ADE*, instead diminish, becoming new triangle *FHE*.¹⁵⁷ Using a stick instead would produce a leftward shift (not illustrated), with the opposite impact on the two triangles.

¹⁵⁵ See Sections III.A and III.B.

¹⁵⁶ For simplicity, I graph the cost-curve shift as though it were identical in magnitude for all producers. Since we hypothesize that naïve producers will pay tax on more units than sophisticates will, that is probably inaccurate; naïve producers will shift more. My general point is unaffected by this simplification.

¹⁵⁷ Lost government revenues from sophisticates, which bear a relationship to the length of line segment *DA*, also grow, as lost revenues now are dependent on the length of the longer segment *HF*. But, at the same time, revenues that would be produced by leaving

Figure 5 can also represent the effect of a stick on reduction of the typical negative externality. Think of households facing a carbon tax. If they are poorer as a result of the tax, they demand less energy, a normal good. Therefore at any given price for a unit of carbon, the household consumes less carbon than in the absence of the tax.

The normative implication here is that the income effect makes gross-up more appealing for many combinations of externality and instrument. By shifting the cost curve rightwards, the income effect of charitable contribution plus carrot, or of carbon tax plus stick, reduces the utility lost from increasing the tax rate sophisticateds perceive (from triangle *ADE* to *FHE* in Figure 5). In contrast, if government did not gross up tax rates, and allowed some naïve producers to perceive a sub-optimal tax rate, the income effects would worsen the outcome (from triangle *ABC* to *FBG* in Figure 5). Full gross-up may still not be ideal, but a partial gross-up may be closer to $1/\theta$ than it would be without income effects. Of course, the opposite is true if government targets negative externalities with carrots or targets positive externalities with sticks. Government should not gross up a partially salient “cash for clunkers” program.

To put this point more intuitively, gross-up magnifies the effects of matching instruments with externalities. The bigger the tax or subsidy, the larger the income effect. If the income effect is working in the opposite direction from the substitution effect—carrots and negative externalities, say—the result will tend to be even less efficient if it is grossed up.

2. *Firms and Output Effects*

Does any of this analysis apply if the externality producers are firms rather than people? It might. Firms are governed by humans,¹⁵⁸ and modern for-profit firms usually pay their managers with a share of firm profits.¹⁵⁹ Taxes and subsidies that affect firm profits therefore may trigger behavioral responses from managers.¹⁶⁰

the tax rate at α (that is, the alternative to gross-up) also drop, because they are related to the length of segment *AC*, which would shift to the shorter segment *FG*. So from a revenue perspective it is ambiguous (without knowing more parameters) whether the shift affects government's choice between the two strategies.

¹⁵⁸ But see *Terminator 2: Judgment Day* (Carolco Pictures, 1991).

¹⁵⁹ See e.g., David I. Walker, *The Law and Economics of Executive Compensation: Theory and Evidence*, in *Research Handbook on the Economics of Corporate Law* 232, 236-37 (Claire A. Hill & Brett H. McDonnell eds., 2012).

¹⁶⁰ For a review of the arguments that managers can respond irrationally to incentive-based pay, see Kelli A. Alces & Brian D. Galle, *The False Promise of Risk-Reducing Incentive Pay: Evidence from Executive Pensions and Deferred Compensation*, 38 *J. Corp. L.* 53, 72-79 (2012).

More straightforwardly, the total output of firms in an industry usually rises and falls with government subsidies or taxation.¹⁶¹ This tends to multiply the impact of any price-instrument policy. Though output effects do not alter the producer's marginal decision about the externality-producing good, they do change the number of producers or number of units produced at the margin. As with income effects, these output effects may affect the trade-offs policymakers face. And again, gross-up will magnify any price instrument's impact on firm output.

Salience could also change the output effect. Over time, output expands in subsidized industries because the government support attracts new investment.¹⁶² If entrepreneurs make investment decisions based on perceived costs and benefits, then output will expand less than expected in the presence of hidden carrots and expand faster than expected with hidden sticks. We should also expect a change in the ratio of naïve to sophisticated producers. For instance, if naïve producers are more attracted than their sophisticated comrades to industries subject to hidden sticks, over time the population of producers will comprise a greater portion of naïves.

It might be argued at this point that salience is largely irrelevant for firms, for whom market competition should drive low- θ actors out of business. This presumes, among other things, that there is enough "smart" money to arbitrage, or bet against, the misguided actors.¹⁶³ In fact evidence suggests a substantial degree of irrational behavior in firms. Most relevant for my claims here, many studies now find significant evidence of overconfidence and other biases among entrepreneurs and other investors.¹⁶⁴ More generally, firm managers and their monitors can never know everything; scholars of managerial behavior find that "satisficing" and other mental shortcuts are pervasive in industry.¹⁶⁵

¹⁶¹ Sterner, note 4, at 167-70.

¹⁶² *Id.* at 167-68.

¹⁶³ For an overview of this assumption, see Gregory La Blanc & Jeffrey J. Rachlinski, In Praise of Investor Irrationality, in *The Law & Economics of Irrational Behavior* 542, 549-57 (Francesco Parisi & Vernon L. Smith eds., 2005).

¹⁶⁴ See e.g., Malcolm Baker, Richard S. Ruback & Jeffrey Wurgler, Behavioral Corporate Finance, in 1 *Handbook in Corporate Finance, Empirical Corporate Finance* 145, 171-76 (B. Espen Eckbo ed., 2007) (reporting other studies finding use of simplified financial assumptions and simplistic targets for borrowing and dividend payouts); Colin Camerer & Dan Lovallo, Overconfidence and Excess Entry: An Experimental Approach, 89 *Am. Econ. Rev.* 306, 314-15 (1999); Ulrike Malmendier & Geoffrey Tate, Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction, 89 *J. Fin. Econ.* 20, 42 (2008).

¹⁶⁵ See, e.g., Sydney Finkelstein, Donald C. Hambrick & Bert Cannella, Strategic Leadership: Theory and Research on Executives, Top Management Teams, and Boards 43-44 (2009); Raphael Amit & Paul J.H. Schoemaker, Strategic Assets and Organizational Rent,

D. *HyperSalience*

Now flip the salience lever from low to high. Just as full salience for sticks initially seemed intuitively appealing, it might at first glance seem as though hypersalient sticks would be especially desirable. After all, hypersalience makes the stick even more effective, with less need for government-imposed penalties. But the problem with a hypersalient stick is that it entails swapping out real money for imagined money. Imagined money might motivate externality producers, but it does not pay the government's bills. I argue, though, that in some unusual instances hypersalient sticks may be a decent second-best option when full-priced sticks are politically unavailable, and that they may also be useful for producing positive externalities.

1. *Sticks*

Again, it is the importance of the stick as a source of revenue that often would make increased salience undesirable. Suppose policy-makers did not alter the price of the hypersalient stick. If price is already set at marginal social damage, the increased perceived cost of the stick will result in over-deterrence. As I showed earlier, it is possible that over-deterrence could be welfare-increasing—if the government were actually collecting revenues from the overcharge.¹⁶⁶ But hypersalience simply produces over-deterrence without bringing in any new revenues. Indeed, because more producers reduce production and so avoid the tax, hypersalience curtails revenues.

Of course, officials could also “gross-down,” or adjust the stick's price downwards to maintain a steady level of deterrence, but that does not solve the revenue problem. Once more, the deterrence-neutral price is simply $1/\theta$. When $\theta > 1$, the new price is a fraction of the old, reducing the revenue collected from each noncompliant producer. This loss is not offset by the addition of new taxpayers: Since subjective deterrence among naïves is unchanged, the number of noncompliant producers also is unchanged, and therefore revenues decline. Moreover, any sophisticates would be under-deterred. Partially grossing down may potentially diminish the overall loss, but there are still few gains to be had.¹⁶⁷

Overall hypersalience looks like it typically will be inferior to other options, but it might still be useful as a second-best choice in some

14 Strategic Mgmt. J. 33, 40 (1993); Sucheta Nadkarni & Pamela S. Barr, Environmental Context, Managerial Cognition, and Strategic Action: An Integrated View, 29 Strategic Mgmt. J. 1395, 1397 (2008).

¹⁶⁶ See Subsection III.A.1.

¹⁶⁷ The net utility of partial gross-down, relative to a full-salience equilibrium, can be expressed as:

situations. Most obviously, if the price of a stick is constrained to be below optimal, but that constraint does not similarly bind efforts to affect the salience of the stick, hypersalience may produce results closer to optimal than the full-salience, too-cheap stick. This scenario seems most likely when the implementation and promotion of the stick are carried out by separate governmental entities, or when there are private third-party beneficiaries. For instance, lobbying might limit Congress' ability to accurately price penalties for nonrenewable energy sources. But advertising by solar- and wind-energy manufacturers might hype the penalty without emphasizing its exceptions. Similarly, one suspects that once the Affordable Care Act's minimum coverage provision is in place,¹⁶⁸ insurers are unlikely to emphasize its exceptions for low-income individuals when advising prospective enrollees.

Hypersalient sticks may also have some added appeal when used to produce positive externalities (or to deter negative externalities attached to inferior goods). Again, in those settings the income effect of the stick is working against its desired substitution effect. Hypersalience and gross-down allow the real tax price to fall, reducing the unwanted income effect. Of course, if the income effect of the stick is also hypersalient, this result is not particularly helpful, so this option would matter mostly when income effects are less salient than substitution effects.

2. *Carrots*

In contrast, hypersalient carrots may well be superior to their only ordinarily-visible cousins.¹⁶⁹ The hypersalient carrot is almost a mir-

$$\begin{aligned}
 U = & -P \left(\frac{1-x}{\theta} \tau w(D_c - D_r) \right) - P_s \left(\frac{x\tau}{\beta_c} \right) \left(\frac{1-x}{\theta} \tau \right) (D_c - D_r) \\
 & - P_n \left(\left(\frac{\tau - \tau x - \theta \tau}{\theta} \right) / \beta_c \right) (D_c - D_r) \tau \\
 & - P_n \left(\frac{\tau^2}{2\beta_c} \left(\left(\frac{\beta_d}{\beta_c} \right) \left(\frac{x^2 - 2x + 1}{\theta^2} - \frac{2 - 2x}{\theta} + 1 \right) + \frac{x^2 - 2x + 1}{\theta^2} - \frac{2 - 2x}{\theta} + 1 \right) \right) \\
 & - P_s \left(\frac{x\tau}{2\beta_c} \left(\frac{\beta_d}{\beta_c} x\tau + x\tau \right) \right).
 \end{aligned}$$

Notably, all of the terms in the equation are negative. The first three reflect lesser revenues: from the tax collected from all producers; from the fact that, when tax rates are lowered, there are fewer sophisticated producers who pay tax; and from naïves who mitigate rather than pay what they perceive as a higher tax. The last two terms are the dead-weight losses from partial over-deterrence of naïves and partial under-deterrence of sophisticates, respectively.

¹⁶⁸ IRC § 5000A.

¹⁶⁹ See Faulhaber, note 14, at 1335, 1340 (suggesting that hypersalient tax expenditures may generate more desirable behavior per dollar forgone by the government).

ror image of the low-salience stick.¹⁷⁰ That is, hypersalient carrots tend towards efficiency because, if combined with offsetting price adjustments, they save government dollars. If the marginal social benefit of another unit of a good is \$100, and producers have $\theta = 2$, the government can spend just \$50 per unit to encourage its production by the naïve. The government can also spend less on sophisticates, who are less motivated by the diminished subsidy, though of course this also means that sophisticates under-produce the beneficial good.¹⁷¹

As I argued earlier, it is likely that the ideal government policy is not to adjust tax to fully offset hypersalience, but instead to only partially gross-down rates.¹⁷² Since producers probably are heterogeneous in their perception of the carrot, the optimal price/salience combination must balance the costs of under-incentivizing sophisticates against any revenue or other savings. It therefore would usually be preferable for any given average θ to set the carrot's price at above marginal social benefit/ θ . If salience is manipulable by the government, some combination of less-than-maximum possible hypersalience and incomplete gross-down may be better than other options.

In addition, the income effects of the carrot might be only normally salient. If so, then grossing-down the carrot's price will naturally reduce any accompanying income effects, and therefore tend to diminish its effectiveness in stimulating most kinds of positive externalities. But tamping down the income effect is actually desirable if carrots are offered to curtail negative externalities or to encourage the production of inferior goods.

Finally on the hypersalience front, many modern carrots, ranging from deductions for charitable contributions to bonus depreciation for

¹⁷⁰ In particular, one can model the welfare effects of the hypersalient carrot as:

$$U = P \left(\frac{1-x}{\theta} \tau w(D_r) \right) - P_n \left(\frac{x\tau}{\beta_c} \right) \left(\frac{1-x}{\theta} \tau \right) (D_r) \\ - P_s \left(\frac{\tau^2}{2\beta_c} \left(\left(\frac{\beta_d}{\beta_c} \right) \left(\frac{x^2 - 2x + 1}{\theta^2} - \frac{2 - 2x}{\theta} + 1 \right) + \frac{x^2 - 2x + 1}{\theta^2} - \frac{2 - 2x}{\theta} + 1 \right) \right) \\ + P_s \left(\left(\frac{\tau - \tau x - \theta\tau}{\theta} \right) / \beta_c \right) (D_r) \tau - P_n \left(\frac{x\tau}{2\beta_c} \left(\frac{\beta_d}{\beta_c} x\tau + x\tau \right) \right).$$

This equation is very similar to the one presented in footnote 127. In this equation, the first two terms represent, respectively, the revenue effects of lowering the subsidy rate and expanding the number of naïve claimants of the subsidy. The next term captures the losses from under-incentivizing sophisticates, and the fourth term captures revenues gained from sophisticates who are not motivated to produce the good under the lower subsidy rate. The final term captures any residual deadweight loss from naïve producers who might be over-incentivized if the carrot is not grossed-down fully.

¹⁷¹ In theory, government could correct the over-incentives problem by enacting a non-linear tax whose rate varied depending on the sophistication of the taxpayer. But such a tax would face a number of obvious practical difficulties in implementation.

¹⁷² See text accompanying notes 121-29.

domestic industrial productions, are delivered through the Code.¹⁷³ Most of these so-called “tax expenditures” operate by reducing taxable income, not the total tax due.¹⁷⁴ Therefore in order to make the value of these carrots hypersalient to taxpayers, someone will have to make clear the beneficiary’s marginal tax rate.¹⁷⁵ But that could have other unexpected welfare effects. For example, if taxpayers are sensitive to the marginal tax rate when they decide how much labor effort to expend, magnifying the salience of individuals’ marginal rates could distort work decisions. The impact of this distortion could well exceed any benefits from the carrot.

E. Correlations

Lastly, I have assumed until now that naïves and sophisticates are randomly distributed across the population. If being a sophisticate is instead correlated with some other factors, the analysis could change.¹⁷⁶ To offer three examples, sophistication could be correlated with the costs of reducing negative externalities (or producing positive ones), with sensitivity to price instruments, and with the efficiency of the price instrument for that producer.

Take first a correlation between sophistication and costs. Suppose, for example, that sophisticates are also better-informed about new technologies and therefore have already adopted production methods that would allow them to adapt to a low-emissions regime. In that scenario, for somewhat technical reasons recited in the margin, it is more plausible that full gross-up could be more appealing than partial gross-up.¹⁷⁷ More interesting are correlations between salience and a

¹⁷³ For an overview of these provisions, see Joint Comm. on Tax’n, 113th Cong., Estimate of Federal Tax Expenditures for Fiscal Years 2012-2017, at 2-28 (Comm.Print 2013), available at <http://www.jct.gov/publications.html?func=startdown&id=4503>.

¹⁷⁴ That is, the expenditures largely take the form of a deduction or exclusion, rather than a credit. See Batchelder et al., note 140, at 24.

¹⁷⁵ For evidence that taxpayers may be unaware of their marginal rate, see Jeffrey B. Liebman & Richard J. Zeckhauser, Schmeduling (2004) (unpublished manuscript), available at http://emlab.berkeley.edu/users/webfac/auerbach/e231_sp05/schmeduling.pdf, and for a review of earlier studies, see Steven M. Sheffrin, Perceptions of Fairness in the Crucible of Tax Policy, in Tax Progressivity and Income Inequality 309, 309-34 (Joel Slemrod ed., 1994).

¹⁷⁶ See Allcott et al., note 7, at 79, for more discussion of correlations between cognitive or willpower failures and externality or internality production.

¹⁷⁷ In essence, there then would be two cost curves, one above *tau* and farther to the right for sophisticates, and one below *tau* and farther to the left for naïve producers. Recall that there is a zero lower bound, at the right-hand side of the graph, representing the maximum possible reduction of emissions. See note 128 (discussing this corner solution). Shifting the cost curve rightwards would bring sophisticates closer to this boundary. If losses are bounded enough, there may be little to be gained from reducing the degree of gross-up. In other words, if most producers would eliminate all pollution even with gross-

producer's responsiveness to price instruments in general. In particular, if naiveté ($\theta > 1$) is correlated with lower responsiveness to price instruments, hypersalience would allow the government in effect to set a "nonlinear" tax, that is, to allow for different prices for different portions of the population. To see why this is useful, step back for a moment and suppose that sticks typically are imposed at some time after the negative externality is generated, and that discount rates—the rate at which we compare present to future costs—vary across producers. Extensive evidence now suggests that some humans are excessively impatient—they grab present rewards without fully considering the later consequences.¹⁷⁸

Galle and Utset and Heutel each note that in this case uniform Pigouvian pricing under-deters impatient actors.¹⁷⁹ That is, because the good is consumed now, but the penalty is imposed later, the impatient are not as sensitive to penalties as the government expected. If the government could, it would threaten the impatient with larger sticks. But as most commentators have noted, it is difficult to see how one could easily let prices vary from individual to individual based on characteristics that individuals can control or misrepresent.¹⁸⁰ Therefore the government usually has to trade off fully deterring the impatient against over-detering the normally patient.

Hypersalience could be a useful tool for overcoming this problem. Impatience and the propensity to treat sticks as hypersalient could well be correlated. Theorists of satisficing and other mental shortcuts similar in impact to salience note that impatience could be an explanation for satisficing behavior.¹⁸¹ Impatient people are unwilling to invest the effort now to get information that will help them in the future. That may be unfortunate for these impatient individuals, but notice how handy it is for the government. By making its stick hypersalient, the government increases perceived costs, but this increase is

down, the government might as well get the revenue benefits of full gross-up from other producers.

¹⁷⁸ See generally George Ainslie, *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person* 56-96 (reissue ed. 2010); Shane Frederick, George Loewenstein & Ted O'Donoghue, *Time Discounting and Time Preference: A Critical Review*, in *Advances in Behavioral Economics* 162 (Colin F. Camerer, George Loewenstein & Matthew Rabin eds., 2004); Richard H. Thaler, *Some Empirical Evidence on Dynamic Inconsistency*, in *Quasi Rational Economics* 127 (1991).

¹⁷⁹ Galle & Utset, note 154, at 74-82; Garth Heutel, *Optimal Policy Instruments for Externality-Producing Durable Goods under Time Inconsistency* 3-4 (Nat'l Bureau Econ. Research, Working Paper No. 17083, 2011), available at www.nber.org/papers/W17083.pdf.

¹⁸⁰ See, e.g., Kaplow & Shavell, note 34, at 6 n.7.

¹⁸¹ See Ted O'Donoghue & Matthew Rabin, *Procrastination in Preparing for Retirement*, in *Behavioral Dimensions of Retirement Economics* 125, 125-26 (Henry J. Aaron ed., 1999).

concentrated among the impatient, the group that needs to face a higher price.¹⁸²

A third interesting potential correlation would be a relationship between the likelihood of being a naïve producer and being an inframarginal one. Recall that an inframarginal producer is someone who would have produced the externality even without the government's intervention. Money spent incentivizing these actors, especially costly carrot dollars, is largely wasted.¹⁸³ Suppose salience is correlated with marginality—for example, perhaps those who are uncertain about whether to donate undertake more of an investigation into the benefits of giving. Then when carrots must be claimed, salience functions as a traditional costly screen because it targets incentives to those whose behavior is most dependent on them.¹⁸⁴

Aid to the poor may be an example of where this kind of screening device could be useful.¹⁸⁵ It is well known that considerably fewer than 100% of eligible recipients claim most forms of government support,¹⁸⁶ and salience could be one reason why.¹⁸⁷ Chetty and Saez find evidence suggesting that the complexities of the EITC contribute to taxpayers failing to take full advantage of the Earned Income Tax Credit (EITC).¹⁸⁸ But suppose that highly-educated households are

¹⁸² See Allcott et al., note 7, at 75, for more discussion in the “internality” context.

¹⁸³ See Stefanie Engel, Stefano Pagiola & Sven Wunder, 65 *Ecological Econ.* 663, 670 n.188 (2008).

¹⁸⁴ See, e.g., Allcott et al., note 7, at 82, 84 tbl.3.; see also Jonathan S. Masur, *Costly Screens and Patent Examination*, 2 *J. Legal Analysis* 687, 688-91 (2010), for more discussion of screening techniques.

¹⁸⁵ For arguments in favor of screening beneficiaries of government programs, see Albert L. Nichols & Richard J. Zeckhauser, *Targeting Transfers Through Restrictions on Recipients*, 72 *Am. Econ. Rev.* 372, 372-77 (1982). For skeptical commentary on screening, see David A. Super, *Offering an Invisible Hand: The Rise of the Personal Choice Model for Rationing Public Benefits*, 113 *Yale L.J.* 815, 844-56 (2004).

¹⁸⁶ See e.g., Karl Kronebusch & Brian Elbel, *Enrolling Children in Public Insurance: SCHIP, Medicaid, and State Implementation*, 29 *J. Health, Pol., Pol'y & L.* 451, 452 (2004).

¹⁸⁷ Cf. Pat Redmond, Judith Solomon & Mark Lin, *Ctr. on Budget & Pol'y Priorities*, *Can Incentives for Healthy Behavior Improve Health and Hold Down Medicaid Costs?* 2 (2007), available at <http://www.cbpp.org/files/6-1-07health.pdf> (reporting a study that found potential beneficiaries did not understand financial incentive programs). Prior evidence finds important roles for psychological factors in benefit uptake rates, but these studies mostly focus on participants' attitudes towards government, the program, and their own self-image in relation to those. See e.g., Julie Christian & Dominic Abrams, *A Tale of Two Cities: Predicting Homeless People's Uptake of Outreach Programs in London and New York*, 26 *Basic & Applied Soc. Psychol.* 169, 169-80 (2004); Sean Nicholson-Crotty, *The Impact of Program Design on Enrollment in State Children's Health Insurance Programs*, 35 *Pol'y Stud. J.* 23, 28 (2007) (noting most scholars attribute low uptake rates to “stigma”).

¹⁸⁸ Raj Chetty & Emmanuel Saez, *Teaching the Tax Code: Earnings Responses to an Experiment with EITC Recipients*, 5 *Am. Econ. J. Applied Econ.* 1, 1-2 (2013). Perhaps similarly, a number of studies find that government outreach efforts improves uptake rates, Nicholson-Crotty, note 187, at 29, though this increased enrollment may also be due to the

more likely to understand the EITC's cumbersome limitations rules, and therefore would be more "sophisticated"—the EITC is not hyper-salient for them. These households have higher earning potential on average¹⁸⁹ and so arguably should not be the target of government support. The effect of salience, then, would be that the neediest households would be those most likely to apply for the EITC, increasing its efficacy. Of course, it also seems possible that benefits could be least salient to those who need them most.¹⁹⁰ Government should work to remedy these kinds of negative correlations.

One disclaimer to these possibilities is that I have assumed so far that individuals who are insufficiently aware of a carrot will not receive it. Think here of charitable contribution deductions that must be claimed on a tax return, or the patent application that must be filed, or the pages-long information form for federally-subsidized student loans. Yet some carrots might be claimed automatically. Price supports for milk, corn, and other farm products¹⁹¹ have sometimes been designed as automatic carrots. The government boosts prices by buying up supply during boom years.¹⁹² Farmers need do nothing but literally reap the benefits and bring them to market. Presumably the price supports are in place to encourage steady production of nationally-important domestic food supplies.¹⁹³ Some farmers might never know that the price supports exist, and if they would have farmed anyway, they get the benefits regardless.¹⁹⁴

reduced effort and stigma attached to programs for which government workers try to enroll beneficiaries. See Kelly M. Purtell, Elizabeth T. Gershoff & J. Lawrence Aber, *Low Income Families' Utilization of the Federal "Safety Net": Individual and State-Level Predictors of TANF and Food Stamp Receipt*, 34 *Child Youth Serv. Rev.* 713, 722 (2012).

¹⁸⁹ Gary S. Fields, *Accounting for Income Inequality and its Change: A New Method, with Application to the Distribution of Earnings in the United States*, 22 *Research in Labor Econ.* 1, 29-30 (Soloman W. Polacheck ed., 2003) (reporting that education is the most important predictor of household income).

¹⁹⁰ Purtell et. al., note 188, at 721-23 (reporting that uptake increased with measures of family need, but noting that non-English speaking households had lower uptake rates); Raj Chetty, John N. Friedman & Emmanuel Saez, *Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings*, 103 *Am. Econ. Rev.* 2683, 2684-85 (2013). (describing regional pattern of awareness of EITC provisions).

¹⁹¹ As far as I am aware, there are no carrots for carrots, however.

¹⁹² For a summary of the current milk price-support programs, including both the government buying policy as well as the similarly automatic system of price controls for milk buyers, see Ralph M. Chite, *Cong. Research Serv.*, IB97011, *Dairy Policy Issues 5-8* (2005), available at http://assets.opencrs.com/rpts/IB97011_20051110.pdf.

¹⁹³ Jim Monke, *Cong. Research Serv.*, RL33271, *Farm Commodity Programs: Direct Payments, Counter-Cyclical Payments, and Marketing Loans 1-3* (2006), available at <http://www.nationalaglawcenter.org/wp-content/uploads/assets/crs/RL33271.pdf>.

¹⁹⁴ Inframarginal farmers drive up the expense of cost-stabilization plans by generating excess production the government must buy during boom years. Farm subsidies may also be captured by nonfarmers, such as owners of farmland. Barrett E. Kirwan, *The Incidence*

Accountants, and more recently TurboTax, might supply another example.¹⁹⁵ If this author's experience is representative, even well-informed taxpayers may be surprised at the end of the year to discover some of their available deductions. These tax benefits do have to be claimed, but because the taxpayer does not have to take any action on her own to claim them—except for hiring an accountant or buying tax preparation software that asks, “Did you weatherize your home this year?”—they are effectively automatic. As a result, using the Code to deliver carrots can increase uptake rates, but at the potential cost of increasing payouts to inframarginal recipients.

IV. IMPLICATIONS FOR CHOICE OF INSTRUMENTS

For all the possibilities that the combination of salience and Pigouvian pricing has offered in my analysis so far, I have not yet touched on perhaps the most important implication. As I mentioned earlier, standard economic analysis strongly favors sticks over carrots in most settings, though I have also suggested that that result should be qualified in some circumstances.¹⁹⁶ I now argue that it is possible that salience could increase carrots' appeal, perhaps to the point where carrots are viable policy alternatives to sticks in some situations. If I am right, the triumph of carrots would be a dramatic reversal of well-settled conventional wisdom. But admittedly, because my suggestions here depend on strong assumptions about the way that salience works, these possibilities are even more tentative than those of the previous Part.

A. *Repeated Games*

At least since Coase, the key factor weighing against carrots is the incentives carrots offer for gamesmanship.¹⁹⁷ Paying off today's polluters to clean up their act encourages future generations of entrepreneurs to invest in polluting.¹⁹⁸ Rewarding good deeds with cash indicates to do-gooders that they should wait for payouts instead of volunteering.

of U.S. Agricultural Subsidies on Farmland Rental Rates, 117 J. Pol. Econ. 138, 151-53 (2009).

¹⁹⁵ For further discussion of the relationship between TurboTax and tax administration, see Austan Goolsbee, *The TurboTax Revolution: Can Technology Save Tax Complexity?*, in *The Crisis in Tax Administration* 124, 128-37 (Henry J. Aaron & Joel Slemrod eds., 2004); Zelenak, note 74, at 98-116.

¹⁹⁶ See Galle, note 4, at 813-40.

¹⁹⁷ See Wiener, note 4, at 726.

¹⁹⁸ *Id.* at 726-27.

But suppose that only *current* carrot recipients can perceive the carrot. A highly opaque carrot could reach just these existing producers, but would not encourage new entries into the field. If so, that would help considerably to mitigate carrots' tendency to attract new negative externality producers who want to be paid to stop. For instance, $\theta=1$ could be strongly correlated with being someone who is already producing an externality. Farmers probably follow the details of farmers' aid programs more closely than do architects or archaeologists. Recipients of new carrots are likely to be highly informed about the carrot program's details because they were the ones who lobbied for it, while the rationally ignorant general public paid little attention.¹⁹⁹ Arbitrageurs may in time close this gap, however, which perhaps argues for time-limiting low-salience carrot programs.

Confidential settlement agreements offer a possible real-world example of this phenomenon. In "responsive" or cooperative regulation, government regulators may waive penalties for violations committed by "good" or cooperative actors, especially if those actors voluntarily disclose their misstep to the regulator.²⁰⁰ Often a condition of the waiver or settlement agreement is that the regulator will not announce the terms.²⁰¹ Some commentators criticize this practice, justifiably, for reducing government accountability and perhaps diminishing the expressive component of punishment.²⁰² But it also serves to reduce the salience of the potential carrot for those who have not yet entered the cooperative program. In this way, the regulator may be able to shrink any temptation for private actors to do wrong on the assumption they will still get a carrot (a welcome departure from expected costs) later down the line. While actors may know

¹⁹⁹ Cf. Anthony Downs, *An Economic Theory of Democracy* 240-59 (1957) (setting out theory of rationally ignorant voters). For evidence on the extent of voter information, see Stephen Earl Bennett, *Trends in Americans' Political Information, 1967-1987*, 17 *Am. Pol. Q.* 422, 426-29 (1989) (finding that voter knowledge was consistently low for the twenty years studied).

²⁰⁰ See generally Ian Ayres & John Braithwaite, *Responsive Regulation: Transcending the Deregulation Debate* (1992) (providing a theoretical overview); Jodi L. Short & Michael W. Toffel, *Making Self-Regulation More Than Merely Symbolic: The Critical Role of the Legal Environment*, 55 *Admin. Sci. Q.* 361 (2010) (providing a helpful discussion and quantitative assessment).

²⁰¹ Christopher A. Wray & Robert K. Hur, *Corporate Criminal Prosecution in a Post-Enron World: The Thompson Memo in Theory and Practice*, 43 *Am. Crim. L. Rev.* 1095, 1176 (2006); see also Cristie Ford & David Hess, *Can Corporate Monitorships Improve Corporate Compliance?*, 34 *J. Corp. L.* 679, 702 (2009) (noting that agreements are confidential); Brandon L. Garrett, *Structural Reform Prosecution*, 93 *Va. L. Rev.* 853, 885 (2007). In part, this appears to be motivated by a desire to avoid follow-on litigation by third parties, such as other regulators or private plaintiffs. *Id.* at 1173-74.

²⁰² See Ford & Hess, note 201, at 702, 723; see also Ronald Wright & Marc Miller, *Honesty and Opacity in Charge Bargains*, 55 *Stan. L. Rev.* 1409, 1410-13 (2003) (making these points about opacity of plea bargaining generally).

that the cooperation program exists, their uncertainty concerning its value reduces the resulting moral hazard.

B. Inframarginal Producers

This kind of selective salience can also help overcome another carrot shortcoming in the production of positive externalities. Inframarginality tends to diminish the efficacy of carrots offered to benefit producers. Again, some good-hearted people would generate benefits for others without explicit rewards, but these folk are usually able to collect their carrot regardless. If the producers anticipate that a carrot might be forthcoming (or increasing) in the future, they may hold off on their beneficial activity until it arrives, so that carrots can actually crowd out the externalities they are aimed at creating.²⁰³ But suppose a correlation between high θ and producers who are *not* inframarginal. That is, what if only marginal producers see the carrot? For instance, charitable fundraisers may concentrate their efforts on individuals who would not otherwise give or give as much, increasing awareness of tax benefits among that population.²⁰⁴

In the other direction, hypersalience could also potentially be used to trim back inframarginality. If sophisticates are mostly inframarginal, dropping the price of the carrot by definition will not reduce their desire to produce. Policymakers could then improve welfare by lowering the carrot price and increasing hypersalience. Naïve producers, attracted by the overly-visible subsidy, will continue to contribute, while the sophisticates will not change their behavior.

Why would sophistication and inframarginality be correlated? By definition the inframarginal producer is less inclined to free ride on the efforts of others when it comes to positive externalities tied to the given good. She therefore could also be more willing to invest in lobbying for greater subsidies for that good, or at least to invest in following the issue closely. Less abstractly, inframarginal donors could be motivated by ideological commitment to their cause, and that would make them follow its progress more intensively.²⁰⁵ In the charitable context, inframarginal donors are often motivated by "warm glow," or

²⁰³ See William J. Baumol & Wallace E. Oates, *The Theory of Environmental Policy* 212 (Cambridge Univ. Press 1998) (1975).

²⁰⁴ There is evidence to support that possibility. A number of studies find that fundraising usually increases donations at the individual firm level. See e.g., Adrian Sargeant & Jurgen Kahler, *Returns on Fundraising Expenditures in the Voluntary Sector*, 10 *Nonprofit Mgmt. & Leadership* 5, 10-16 (2003). This result could be consistent with a story in which firms that fundraise more intensively simply shift donors away from other firms, however.

²⁰⁵ See Allcott et al., note 7, at 9-10 (suggesting that "'green' consumers who derive warm glow from purchasing the more energy efficient product . . . might give more weight to energy efficiency. . .").

the good feelings that come with giving.²⁰⁶ Some evidence suggests that part of this warm glow is that being a philanthropist is a status symbol, and that donations are a luxury good that appeals especially strongly to the wealthy who want to signal their wealth.²⁰⁷ Inframarginality would then correlate with sophistication to the extent that sophistication is correlated with wealth.²⁰⁸

Alternately, consider the possibility I raised earlier that salience is an effective form of screen for sorting government benefit recipients.²⁰⁹ If needier families are less attuned to the downsides of a program, making it hypersalient for them, the program overall can be more efficient. These kinds of effects may mitigate the welfare costs of carrots enough to make them worthwhile as policy choices, although perhaps still not as desirable as a comparable stick.

C. Income Effects

A third flaw in most carrot programs involves income effects. Recall that income effects make carrots for positive externalities more effective by increasing demand for the desired good (as long as it is a normal or luxury good).²¹⁰ If this income effect remains somewhat salient even as the substitution effect of the carrot is submerged, government can increase production without triggering any of the unwanted side effects of a dose of carrots. In most contexts the income effect from changing one item in a household's budget will be small enough that this strategy would not likely be cost-effective. But if other policy options are unworkable, the low-salience carrot is worth keeping in the toolkit.

Another possible income effect involves a high- θ of substitution combined with a low- θ of income. In other words, producers know the cost of each item, but are not as aware of its total impact on their budget. If that seems like an unlikely combination, imagine that we know there is a tax on some tempting item—this author is fond of chocolate-covered cashews—but we are not sure how much of it we will be able to keep ourselves from consuming. The evidence is that

²⁰⁶ See Lise Vesterlund, *Why Do People Give?*, in *The Nonprofit Sector: A Research Handbook* 568, 572-81 (Walter W. Powell & Richard Steinberg eds., 2d ed. 2006).

²⁰⁷ See Amihai Glazer & Kai A. Konrad, *A Signaling Explanation for Charity*, 86 *Am. Econ. Rev.* 1019, 1019-621 (1996).

²⁰⁸ But see Goldin & Homonoff, note 62, at 325 (“attentiveness to cigarette register taxes declines monotonically by income”).

²⁰⁹ See text accompanying notes 183-89.

²¹⁰ See Figure 1 and text accompanying note 48.

humans are bad at predicting their own willpower.²¹¹ At the time we draw up our budgets, we might not know how much to set aside for cashew costs even if we know exactly the price of each can.²¹² Manuel Utset and I earlier made this argument about energy taxes.²¹³

In any event, assuming the high- θ of substitution/low- θ of income combination is plausible, it can be somewhat useful to carrot-wielding policymakers. Carrots confound government's goals when aimed at negative externalities, since the increased wealth of recipients makes them produce or demand more of the unwanted good. If the income effects of the carrot are obscured, this impact can be softened, which again could make carrots marginally more useful—or at least mitigate the damage from a flawed political process that overly relies on them.

V. LEARNING & TRANSITIONS

The discussion so far has assumed a mostly static world in which the salience of a tax instrument is relatively fixed. But despite the best efforts of cable news and reality television producers, humans do occasionally move from ignorance to wisdom. If people quickly learn to understand sticks and carrots, salience may be of relatively little importance. Theory and some evidence suggest, though, that salience effects can persist for long periods. That implies, happily, that the work of ploughing through the previous four Parts was not in vain. But lawmakers should also give some thought to how to manage gradual shifts in public understanding of price instruments.

A. *Theory & Evidence on Learning*

I have argued in other work that we should not expect learning to be pervasive.²¹⁴ If individuals “intentionally” ignore hidden prices because they believe the cost of calculating the price exceeds the benefits, the same often will be true of the cost of educating oneself to be more mindful in the future—or the same but even more so, because the benefits of education are postponed farther into the future.²¹⁵

Individuals whose cognitive limitations curtail their ability to account for a cost—for example, because they lack the willpower to

²¹¹ See Stefano DellaVigna, *Psychology and Economics: Evidence from the Field*, 47 J. Econ. Literature 315, 320 (2009); Ted O'Donoghue & Matthew Rabin, *Choice and Procrastination*, 116 Q.J. Econ. 121, 150 (2001).

²¹² See Chetty et al., note 60, at 1145-48, 1174 (explaining the household budgeting process under uncertainty about hidden taxes).

²¹³ Galle & Utset, note 154, at 57-60.

²¹⁴ Galle, note 6, at 89-93. For arguments that learning may be difficult in the context of the charitable contribution deduction, see Faulhaber, note 14, at 1323-26.

²¹⁵ Galle, note 6, at 88.

make themselves engage in the painful act of thinking about math—may similarly be limited in their capacity to force themselves to learn.²¹⁶ Weak incentives and low credibility limit the usefulness of third-party instructors.²¹⁷ For example, competitors of third-party beneficiaries of a hypersalient stick might have weak incentives to disabuse consumers because each competitor would capture only a fraction of the resulting business.²¹⁸ Consumers may know precisely the extent of a tax but lack the willpower or wherewithal to apply that information to purchase decisions.²¹⁹

Even if every economic actor eventually learns every detail of every price instrument, the fact would remain that all sticks are hidden sometimes. Logically, when a new price instrument is introduced, there is inevitably a period in which not all of the relevant economic actors are aware of it.²²⁰ Like any other piece of information that influences commodity prices, understanding of a new stick's legal structure of may diffuse unevenly across the economy.²²¹

While there is some evidence that individuals can learn about hidden pricing, the findings often imply significant gaps in public understanding. For example, in the tax context, taxpayers can be taught to negotiate the complexities of the EITC.²²² But that same study implies that in the absence of professional assistance, most of the study participants were baffled by the EITC. In my work with Jonathan Klick, we suggest that our findings on the effect of the Alternative Minimum Tax on local spending can be explained (among other possible ways) by the willingness of taxpayers who expect AMT liability to invest in learning about its impact.²²³ Other taxpayers who pay AMT do not appear to be as responsive, perhaps reflecting their ignorance of its many nuances.²²⁴

Learning may also take a fair bit of time. Bakija & Heim report that donors to charity initially seem to neglect changes in tax rewards

²¹⁶ See Galle & Utset, note 154, at 76-78.

²¹⁷ See Galle, note 6, at 91-92.

²¹⁸ See Gabaix & Laibson, note 5, at 505-23, 526-32.

²¹⁹ For example, Chetty, Looney, and Kroft report that the consumers they studied knew fairly precisely the sales tax rate; they simply failed to apply that knowledge to purchase decisions. Chetty et al., note 60, at 1165.

²²⁰ For an early theoretical overview of the obstacles to the formation of a full-information equilibrium, see Steve Salop, *Information and Monopolistic Competition*, 66 *Am. Econ. Rev. Papers & Proc.* 240, 240-45 (1976). For later theory, see Dirk Bergemann & Juuso Välimäki, *Market Diffusion with Two-Sided Learning*, 28 *RAND J. Econ.* 773, 776-78 (1997).

²²¹ For evidence, see Atakan Yalçın, *Gradual Information Diffusion and Contrarian Strategies*, 48 *Q. Rev. Econ. & Fin.* 579, 586-97 (2008).

²²² Chetty & Saez, note 188, at 1.

²²³ Galle & Klick, note 14, at 246.

²²⁴ *Id.* at 225-36.

for giving, but that over time, a year or so on average, there are signs that some adapt.²²⁵ Survey data from Goldin and Listokin on the low public awareness of the marginal value of charitable contributions suggests that the transition period may be longer than the average reader may expect.²²⁶ Miller and Mumford find some evidence consistent with learning among higher-income households, but not until four years after the policy change they study.²²⁷ And I have argued that a prominent behavioral public finance puzzle can be explained by learning, with the evidence suggesting that voters take several years to understand the combined effects of federal and state fiscal instruments.²²⁸

Whether learning is durable is another question. Sumit Agarwal, John Driscoll, Xavier Gabaix, and David Laibson find, for example, that financial sophistication first rises but then dwindles with age.²²⁹ Other lab studies report similarly that subjects sometimes recall only a portion of what they learn.²³⁰ Readers who have taken a final exam recently, or graded one, may have some familiarity with the latter phenomenon.

Third-party actors can also shape the accuracy of public perceptions. In some regions of the country there appear to be social networks that facilitate the spread of EITC information, but in others these networks are absent.²³¹ Rupert Sausgruber and Jean-Robert Tyran find peer-to-peer learning about tax-like instruments in a laboratory setting.²³² But there can also be third-party misinformation.²³³ Goldin and Listokin report that a statistically significant portion of

²²⁵ Bakija & Heim, note 146, at 636; see also Bridget Terry Long, *The Impact of Federal Tax Credits for Higher Education Expenses*, in *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It* 101, 122-28 (Caroline M. Hoxby ed., 2004) (reporting results of two studies finding low awareness of education tax credits passed two years earlier, even among families that were eligible).

²²⁶ Goldin & Listokin, note 70, at 7-8.

²²⁷ Miller & Mumford, note 67, at 18-19.

²²⁸ Brian Galle, *Federal Grants, State Decisions*, 88 B.U. L. Rev. 875, 924-30 (2008).

²²⁹ Sumit Agarwal, John C. Driscoll, Xavier Gabaix & David Laibson, *The Age of Reason: Financial Decisions over the Life Cycle and Implications for Regulation*, Brookings Papers on Econ. Activity, Fall 2009, at 51.

²³⁰ See Galle, note 6, at 92-93, for a summary.

²³¹ Chetty et al., note 190, at 2-3.

²³² See generally Rupert Sausgruber & Jean-Robert Tyran, *Are We Taxing Ourselves? How Deliberation and Experience Shape Voting on Taxes*, 95 J. Pub. Econ. 164 (2011). Note, however, that in the Sausgruber and Tyran experiment peers had nothing to gain from the financial ignorance of their fellow participants. In reality, as Gabaix & Laibson, note 5, at 505-23, argue, sophisticated consumers can often benefit from the ignorance of others. For empirical evidence to this effect, see Ryan Bubb & Alex Kaufman, *Consumer Biases and Mutual Ownership*, 105 J. Pub. Econ. 39 (2013).

²³³ Cf. Allcott et al., note 7, at 83 (noting that private firms can hide information the government would like to be more salient).

their survey respondents over-valued the home mortgage interest deduction.²³⁴ This may be due to advertising efforts. Two recent studies find that homeowners who do not see their bi-annual property tax bill, but instead allow the bill to be paid on their behalf by the mortgage lender, are less responsive to their property taxes.²³⁵

B. Learning and Price-Instrument Transitions

At a minimum, then, the lessons of the last Section imply a need for careful treatment of transitions into new price-instrument regimes. After a new price instrument is introduced, it likely has low salience for some fraction of the population. Or, if it has a highly visible “sticker” price but lots of complex exceptions, it may be hypersalient. For some period of time, the policymaker’s tool will not have its expected effects.

The analysis of Part III implies that the efficient price of a price instrument should change as the public’s perception of it changes. For example, if we think that the public will tend to have $\theta < 1$ in the first year of a new stick program, the price of the stick therefore should be grossed-up above its full-salience price. As the stick’s salience increases with time, the price should decline. Sticks, in other words, should transition in downwards.

One immediate word of caution about this result is that it may increase arbitrage opportunities for pre-existing owners of affected property. First, note that when salience varies among the population, there usually will be chances for trades that benefit the better-informed party. For instance, suppose that the salience of a stick is correlated with owning an asset that will be subject to that stick. When the over-priced stick is introduced, existing owners would then be more likely to discount the value of their property by the correct amount, while nonowners would over-value the property. Existing owners could then sell to nonowners at a profit.²³⁶ Because a stick that “transitions down” would necessarily involve an initially bigger

²³⁴ Goldin & Listokin, note 70, at 9.

²³⁵ Cabral & Hoxby, note 82; Hayashi, note 83.

²³⁶ For evidence, see Sebastien Bradley, *Capitalizing on Capped Taxable Values: How Michigan Homebuyers Are Paying for Assessment Limits 3* (Apr. 2011) (unpublished manuscript), available at http://www-personal.umich.edu/~sebbBrad/Sebastien_Bradley_JMP.pdf (reporting that “homebuyers dramatically overcompensate sellers of homes with temporarily low tax obligations as if such obligations would persist indefinitely”). The new owners in turn could potentially pass on the hot potato to new unwitting buyers, although at some point (as standard equilibrium theory predicts) most buyers will likely notice that the “smart money” is selling (even if the naïve buyer does not know why) and will refuse to pay a premium any longer.

stick, it would increase the rewards sophisticated prior owners could extract from naïve buyers.

Most theorists of legal transitions would argue that allowing pre-existing owners to profit from a new stick is undesirable. A sales premium for prior owners is economically equivalent to a legal rule providing compensation to owners for the lost value of their property.²³⁷ As many other commentators have argued, a compensation rule inefficiently encourages owners to develop their property in ways that would generate negative externalities—or, similarly, allows them to under-invest in efforts to produce positive externalities.²³⁸

But this problem may not be serious and could likely be readily mitigated. Unwanted incentive effects would arise only if owners can predict that any future stick will have low enough salience to generate a market for their arbitrage. Even if this were common, potential solutions include an “exit tax” on pre-existing owners or mandatory disclosure of the true price of the stick to prospective buyers.

Transition policy for carrots is a bit different. In the case of low-salience carrots, again the government should “transition down,” with the initial price being higher than optimal under full salience, and then declining as the public learns about the new instrument’s effects. In a nice change of pace for carrot fans, the challenges here are a bit less daunting than for stick transitions. Since naïve purchasers perceive carrots as less valuable than sophisticated existing owners, there is no arbitrage opportunity for the crafty owner. Of course, if sophistication is not strongly correlated with prior ownership, sophisticates might poach high-value property from naïve owners. Even if society is indifferent about the distribution of the carrot’s benefits, this particular transfer is welfare-reducing, since the new owner now will wastefully over-exploit the carrot. But the transition downward limits the size of the poacher’s prize, and again disclosure requirements could help mitigate the problem.

Another possible design choice for carrots is to make them temporary. In Part IV, I suggested that carrots may be more appealing when

²³⁷ See Wiener, note 4, at 726 n.186.

²³⁸ See, e.g., Lawrence Blume & Daniel L. Rubinfeld, Compensation for Takings: An Economic Analysis, 72 Cal. L. Rev. 569, 590-92 (1984); Louis Kaplow, An Economic Analysis of Legal Transitions, 99 Harv. L. Rev. 509, 529-31 (1986).

Some commentators defend compensation on other grounds. For example, the need to raise money to pay for compensation might discourage inefficient regulation, such as might result from rent-seeking by other interest groups. Daniel Shaviro, When Rules Change: An Economic and Political Analysis of Transition Relief and Retroactivity 77-78 (2000). For skeptical treatment of that argument, see Daryl J. Levinson, Empire-Building Government in Constitutional Law, 118 Harv. L. Rev. 915, 968-71 (2005); Daryl J. Levinson, Making Governments Pay: Markets, Politics, and the Allocation of Constitutional Costs, 67 U. Chi. L. Rev. 345, 373-77, 388-93 (2000).

they are hidden from some portion of the population, whether that be inframarginal producers or individuals who might threaten harm if they thought it would prompt government to pay them. Alternately, as I noted in Part III, low-salience carrots are less harmful than fully salient, inefficient subsidies. Many current tax breaks are obscure and difficult for the general public to notice, but that may be just as well: If more people knew about them, they would claim them.²³⁹ If carrots are only temporarily hidden, these benefits could fade. A possible response, then, would be either to switch from the carrot to a stick as the program ages, or instead simply to limit the time in which the incentive is offered. A number of existing tax incentives have this kind of “sunset” feature.²⁴⁰ That is typically due to the idiosyncrasy of congressional budgeting rules.²⁴¹ My analysis here offers another potential reason to favor *temporary* tax breaks, at least assuming that the tax breaks are to be enacted anyway.²⁴²

Finally, it might be argued that learning should be met, not with transition policy, but instead with an abandonment of cognitively complex price instruments. Learning itself may be costly. Arguably the time and effort invested in analyzing a new price instrument is dead-weight loss for each actor.²⁴³ If externality producers were perfectly rational, they would invest effort in learning up to the point at which further investments no longer produce gains.²⁴⁴ Even producers who chose not to learn would still incur some decision costs: the costs of deciding whether it is worthwhile to learn about the stick.²⁴⁵ In order to hide large prices from the public, the government would need to make those prices very difficult to identify. The implication is that the

²³⁹ Of course, the obscurity of the breaks may also factor into whether they are enacted at all; it may be that low market salience and high political salience is the best combination for these kinds of inefficient giveaways.

²⁴⁰ See Rebecca M. Kysar, *Lasting Legislation*, 159 U. Pa. L. Rev. 1007, 1017-21 (2011).

²⁴¹ See William G. Gale & Peter R. Orszag, *Sunsets in the Tax Code*, 99 Tax Notes 1553, 1553 (June 9, 2003); George K. Yin, *Temporary-Effect Legislation, Political Accountability, and Fiscal Restraint*, 84 N.Y.U. L. Rev. 174, 188-92 (2009). For a normative defense of temporary tax rules, see Yin, *supra*, at 232-53; see also Jacob E. Gersen, *Temporary Legislation*, 74 U. Chi. L. Rev. 247, 266-72 (2007) (offering qualified defense of temporary law generally). For a critique, see Kysar, note 240, at 1051-65.

²⁴² Admittedly, however, the incentives of actors who are aware of the temporary nature of legislation greatly complicate efforts to accurately price temporary price instruments. Agents may shift future activity forward in time to the period of the subsidy, or may ignore the subsidy because obtaining it requires high fixed costs that will not be paid off over the temporary life of the instrument. Gersen, note 241, at 278.

²⁴³ See Kysar, note 240, at 1064 (pointing out that temporary legislation increases planning costs for affected private interests).

²⁴⁴ See John Conlisk, *Why Bounded Rationality?*, 34 J. Econ. Literature 669, 686-87 (1996) (examining incentives for actors to invest in costly learning).

²⁴⁵ See *id.* (describing this “regress” problem).

deadweight loss of learning from a pricey and highly opaque stick could be substantial.

But that story is a worst-case scenario. Low-salience sticks are more promising if low salience is often the product of impatience and procrastination. That is, for some actors it might easily be worthwhile (in the abstract) to analyze a price instrument, but in the moment they face that decision they overly weight the present inconvenience relative to the far-off gains of making a correct decision. In that case, the government can generate meaningful salience effects even though the mental effort needed to fully identify the price effect is small, implying that the deadweight loss of learning is minor.²⁴⁶ Further, the price of learning is likely not purely deadweight loss.²⁴⁷ As your high-school math teacher would argue, thinking about finance helps to develop cognitive skills that are generally useful and builds information that can be shared with others. For example, skill at identifying prices with low market salience may also facilitate recognition of carrots with low political salience, which can constrain rent-seeking by other externality producers. Learning to overcome salience, in other words, can produce positive externalities.²⁴⁸

VI. CONCLUSION

The study of policies intended to sway real humans, rather than to affect the platonic ideal of rational actors, is only beginning. At this stage in our understanding of how individuals cope with complex systems of price and other information, it is too early to say with confidence that a particular approach is clearly right or wrong. My effort here has been to begin to explore some possibilities, given what we know now.

What I think I have shown is that some well-settled assumptions may be incorrect, and that some recent advances can be further refined. Carrots are not necessarily always inferior to sticks, once salience and its potential correlations are taken into account. Fully

²⁴⁶ Colin Camerer, Samuel Issacharoff, George Loewenstein, Ted O'Donoghue & Matthew Rabin, *Regulation for Conservatives: Behavioral Economics and the Case for "Asymmetric Paternalism,"* 151 U. Pa. L. Rev. 1211, 1219 (2003).

²⁴⁷ See Raj Chetty, John N. Friedman, Soren Leth-Petersen, Torben Heien Nielsen & Tore Olsen, *Active vs. Passive Decisions and Crowdout in Retirement Savings Accounts: Evidence from Denmark* (Nat'l Bureau Econ. Research, Working Paper No. 18565, 2012) available at obs.rc.fas.harvard.edu/chetty/crowdout.pdf (reporting beneficial effects of economics training on ability to plan retirement savings).

²⁴⁸ The extent of the positive externality likely depends on legal rules governing the enactment and interpretation of the stick. In particular, if legal rules constrain sophistication from lobbying for special benefits for themselves, their lobbying efforts are more likely to affect others (in some cases positively, in some not).

salient price instruments may not always be ideal. Government should likely adjust the prices of its taxes and subsidies at least to account for the process of learning and diffusion of information across the economy. And these adjustments may be closer to optimal if they balance revenue and other needs against the correction of externalities.

My hope is that these tentative thoughts can be the building blocks for further efforts both theoretical and empirical. For example, if I am on the right track about the relevance of salience for the choice of price instruments, it will be important to have more and better empirics on the relationship between salience and such factors as the propensity to emit externalities. Designing the best price for an instrument will require better data on the distribution of salience across the population, the extent to which it is correlated with wealth and income, and the likelihood that increased prices will themselves affect incentives to learn more about government's regulatory tools. The work, in short, is just beginning.

